DMSTTIAC

Defense Modeling, Simulation and Tactical Technology
Information Analysis Center

DMSTTIAC SOAR 95-01

De Physica Belli

An Introduction to Lanchestrial Attrition Mechanics
Part One

Bruce W. Fowler, Ph.D.
Technical Director and Deputy
Advanced Systems Concepts Office
Research, Development and Engineering Center
U.S. Army Missile Command

Published by:
DMSTTIAC
IIT Research Institute
7501 S. Memorial Parkway, Suite 104
Huntsville, AL 35802

U.S. Army
Missile Command

November 1995

OCT 10 '96 11:07AM ASCO P.1
10/09/1996 13:04 312-567-4889 IIT RESEARCH DEPT G PAGE 01

Telecopy Message

Date: 9 Oct 96

Transmitting Cover Sheet and Pages ____

TO: Renee Gorgus
DMSTTIAC

PHONE: 205.880-0884

FAX:

FROM: Gerri A Koclanis
IIT Research Institute/GACIAC
10 West 35th Street
Chicago, Illinois 60616-3799
Phone: 312.567.4587
Fax: 312.567.4889

Guidance and Control
GACIAC
Information Analysis Center

GACIAC is operated by IIT Research Institute Chicago, Illinois since 1977 under sponsorship of DTIC

REMARKS:

Renee: Per telephone conversation today, DMSTTIAC SOAR 95-01 De Physica Belli has missing pages. I received a phone call today from Pat Mawby, at DTIC and she and I looked at the report to verify the following:

SECTION XVII Pages 4 and 6 are missing

SECTION XVIII Page 6 missing

SECTION XX Pages 9, 15, 17, 19, 23 and 26 missing.

Appendix E pages are not numbered along with several blank pages and figures in the document.

After speaking to you, Dr Fowler was aware of these pages and in essence they are not missing but pagination is in error. Please fax me a statement regarding the pages so that I can send it to DTIC.

Thank you,

Gerri

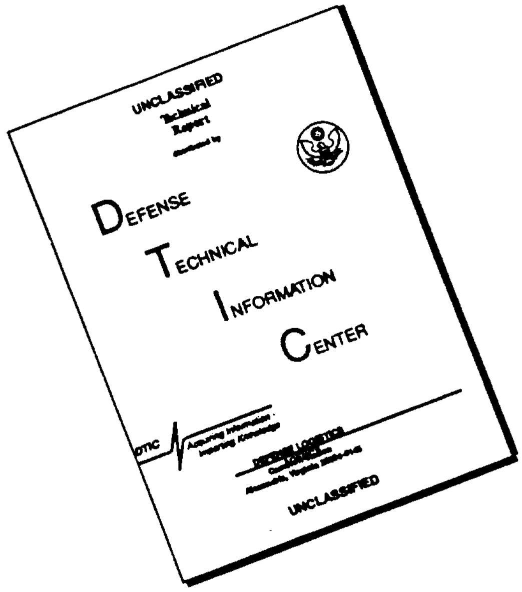
FAX # 880-0886
Phone - 880-0884

I know of course
Emma Fowler

OCT-09-96 WED 15:54 IIT RESEARCH INST FAX NO. 2058800886 P.01/01

OCT-24-96 THU 09:46 IIT RESEARCH INST FAX NO. 2058800886 P.01/01

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

NOTICES

State of the Art Review. This state-of-the-art review has been published by the Defense Modeling and Simulation Tactical Technology Information Analysis Center (DMSTTIAC) as a service to both defense and non-defense agencies, academia and industry. DMSTTIAC is a DoD Information Analysis Center administered by the Defense Technical Information Center and operated by IIT Research Institute under contract DAAH01-95-C-0310. DMSTTIAC is funded by the Defense Technical Information Center (DTIC) and the Defense Modeling and Simulation Office (DMSO). The Director of DMSTTIAC is Mr. Hunter Chockley. The Contracting Officer is Mr. Scott Savory, Defense Electronics Supply Center (DESC), Dayton, Ohio. The Technical Monitor is Mr. Chalmer D. George, and the Alternate is Mr. Howard C. Race, AMC-Smart Weapons Management Office (SWMO), Attn: AMSMI-SW, Redstone Arsenal, Alabama 35898-5222.

Reproduction and Handling. Unlimited Distribution

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
November 1995

3. REPORT TYPE AND DATES COVERED
State-of-the-Art-Review; November 16, 1995

4. TITLE AND SUBTITLE
De Physica Belli: An Introduction to Lanchestrian Attrition Mechanics, Part I

5. FUNDING NUMBERS
DAAH01-95-C-0310

6. AUTHOR(S)
Bruce W. Fowler, PhD.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
IIT Research Institute/DMSTTIAC
10 West 35th Street
Chicago, IL 60616-3799

8. PERFORMING ORGANIZATION REPORT NUMBER
DMSTTIAC SOAR 95-01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Missile Command
ATTN: AMSMI-SW
Redstone Arsenal, AL 35898-5222

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
This document is available only from DMSTTIAC, IIT Research Institute, 10 West 35th Street, Chicago, IL 60616-3799.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document available only from DMSTTIAC, IIT Research Institute. Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Physics and War are inextricable. While it is often recognized that war is often the generator of progress in physics, the role of physics in war is less often recognized. This work is dedicated to examination of the physics inherent in some of the processes of war. The framework for this examination is the Lanchester model of attrition. Part I of this work is primarily concerned with the basic foundation of this model . This volume, which comprises Part I, presents the mathematical and assumptive basis of homogeneous Lanchester theory, reviews the primary contributors - Lanchester, Chase, Osipov, and Fiske - to the basic theory, and the two basic alternative theories of the same genre. The historical basis for validity of Lanchester theory is reviewed and analyzed. The relationship between Lanchester theory and higher level simulation is introduced along with prefatory expositions of subsequent parts of the work. Part II (to be published) will cover the basic physics inherent in the attrition rate coefficients; Part III will cover heterogeneous Lanchester theory and aggregation; and Part IV will cover special and advanced topics.

14. SUBJECT TERMS Attrition Theory, Lanchester, Lanchester Attrition Theory, Quadratic Lanchester Equations, Mixed Lanchester Equations, Attrition Differential Equations, Quantified Judgement Model, Epstein Model, Osipov's 3/2 Law, Constant Rate Attrition, Exponential Attrition, Helmbold's Attrition, Time Dependent Attrition, Range Dependent Attrition, Stochastic Duels, Stochastic Attrition, Stochastic Lanchester, Linear Attrition, Attrition Coefficients, Attrition Rates, One-on-one, Force-on-Force

15. NUMBER OF PAGES
584

16. PRICE CODE
$23.00

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unclassified

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

DE PHYSICA BELLI

An Introduction to Lanchestrian Attrition Mechanics

Bruce W. Fowler, Ph.D.
Technical Director and Deputy
Advanced Systems Concepts Office
Research, Development, and Engineering Center
U. S. Army Missile Command

and

Adjunct Associate Professor
Department of Physics
University of Alabama in Huntsville

16 November 1995

Foreword

Warfare is as old as recorded human history. War has been especially prevalent in the last 500 years with the increasing conflict between large nation states. A great amount of analysis and thought has been given to the "Art of War". Nine principles of War have been defined: Objective, Offensive, Mass, Economy of Force, Maneuver, Unity of Command, Security, Surprise, and Simplicity. Despite these accepted principles, the science of war has remained elusive. Since World War II, investigators have searched for a theory on the physics of war--"De Physica Belli". Efforts have been more successful with the prominent rise of Operations Research as an analysis tool to assist combat operations. Dr. Bruce W. Fowler uses these modern analytical tools to seek the answer to the following question in this report--"Is there any scientific basis to describe the physics of war?" This report provides the answer to this question. His approach to a physics of war is the application of Lanchestrian attrition mechanics which first appeared in theory in the early 1900's.

Dr. Fowler introduces Lanchester's work and then examines whether Lanchester really was the "father of attrition theory" and the resulting force ratios and attrition coefficients. Lanchester initially claimed that improved tactics, training, doctrine, and morale were not amenable to mathematical analysis. Once the reader generally understands Lanchester's Differential Equations and their solutions, Dr. Fowler proceeds to introduce variations on a theme by carrying Attrition Theory forward until the late 1980s. Some of the topics covered are: stochastic versus deterministic representations; homogeneous versus non-homogeneous forces; dependencies of attrition and attrition rates on time and range, not just on force strength; aggregation and disaggregation; Quantified Judgment Models; Bonder-Farrell Attrition Theory and Ancker-Gafarian Attrition Theory.

"De Physica Belli" is intended to be a general reference and introduction to attrition theory suitable for the combat soldier, the student-soldier, or the military analyst. The manuscript succeeds in that respect and provides a good overall summary of the state-of-practice in attrition theory through 1990. However, given the great advances in modeling, simulation and computational power since 1990, it would not be surprising to see future updates to this work. The mathematical tools of complexity theory, fractal dimensions, fuzzy logic, information theory and the power of scientific visualization of data in interactive computer simulations may offer new and exciting insights into the physics of war. These new developments will most certainly provide opportunities to conduct experiments in the science of warfare that go beyond the limitations inherent in the analysis of historical data.

Part I:

Homogeneous Lanchester Attrition Theory, History, Other Attrition Theories, and Stuff with pH > 7

THIS PAGE INTENTIONALLY BLANK

Table of Contents

Table of Contents
(continued)

Table of Contents
(continued)

Table of Contents
(continued)

Table of Contents
(continued)

Table of Contents
(continued)

PREFACE

This is a book about war.

It is the result of several years interest in the modeling and simulation of warfare. It neither praises nor decries war: war is a social activity of mankind and as such can be avoided or denied no more than any other interaction of man with others of his kind. Clausewitz tells us that war is an extension of policy by other means, while Mao tells us that power grows from the end of a gun. Clearly then, war is social and at best is governed by such social rules as the participants are willing and able to apply.

Why should a physicist write a book about war? The answer has two parts. A physicist, more than most of humanity, looks at the world around him and continually asks "Why?" He applies logic, patience, stubborn determination, and mathematics to the question.

That "Why?" question brought me to the subject of war, and continues to lead me through investigations and studies of it. It also led to writing this book so that others could ask that question with greater efficiency by using what little stubbornness I have been able to apply.

Man has apparently practiced war as long as he has existed. The tool making tradition/development of man is clear. While the application of early chipped rock tools such as choppers and hand axes to warfare can be questioned, the question arises not from the likelihood of their application, but to the nature of warfare in that social environment. Warfare today is viewed as being national in scope (even civil war) and reflecting some cultural conflict (which itself raises the question of how warfare can exist without the benefit of agriculture.) In neolithic times, nations as such did not exist, but familial and tribal level social groups most likely did, and conflict between such groups probably had all of the cultural aspects of war as we think of it. The earliest evidence of warfare as conflict between two (or more) collected forces is found in Neolithic cave paintings[1]. Most Historians neglect warfare prior to the Macedonian Juggernaut of Philip and Alexander, although we now have evidence that the social development of war, its institutions and mechanism is fundamentally older.[2] This is partly due to the lack of recorded history and partly because Macedonia (under Philip) was the first western nation. However, as Jones[3] notes, the primary reason was the emergence, with the Macedonian nation, of the Macedonian army as the first balanced, combined arms army.

The Romans made no strong distinction between technical knowledge and its military applications. Neither, apparently, do the Russians, America's overt rival (until recently?) for dominance of Civilization. In America, we practice an oscillating love-hate relationship with things military. For several years now, this country has practiced an academic apathy for matters of warfare. To this end, there are almost no avenues

for disseminating information on warfare research.

This book is not primarily just for the soldier. By that I do not mean that the soldier should not read it, but I recognize that the profession of soldier is not given to mathematics and analysis. It is however, much given to the rigorous study of stubbornness and patience, to the art of concerted action and deliberate inaction. To the average soldier, a book of analysis would be a punishment no matter how couched. To the occasional soldier with a bent for mathematics and analysis, it would be of insightful use. If I have done my task well, it may even be of abiding and delightful value.

In this book, we limit ourselves (primarily) to some of the aspects of formal war. Formal war is a term that distinguishes warfare characterized by the use on both sides of trained troops under discipline with a rigorous chain of command and a set of formalized goals. Informal warfare such as riots, civic disturbances, terrorism, inquisition, and other spontaneously constituted conflict are thus excluded. (The special case of guerilla conflict is somewhat of a grey area and we shall treat some of the combat aspects of such conflict.) The scope of this book is limited to treating some of the aspects of formal war.

Of particular concern will be the tactical level of formal warfare (or just warfare, as shall hereafter be used synonymously.) The strategic or (recently rediscovered in this country) operational levels of warfare will be devoted little attention. This limitation is dictated not solely by desire but by the fact that the tactical level of warfare is most strongly associated with attrition and attrition is the part of warfare that has been examined most deeply.

It must be noted that the practice of war is an art. However, art has its technical aspects. Just as painting is an art form, it too has its technical aspects - the optical and material technology associated with perspective, color, the functioning of the human eye, the production of paint and canvas. Similarly, the art of war has technical components that support its execution. This book deals with some of those technical components.

This does not mean that this book is intended to have an audience of soldiers. As the execution and appreciation of painting cannot be totally technical, the execution and appreciation of warfare cannot be totally technical as well. But in both instances, there are technical factors and contributions to both the execution and appreciation of painting and warfare. The painter cannot successfully practice his art without knowledge and use of the technical aspects of his tools and methods. Neither can the manufacturer of art supplies be ignorant of the technical aspects of painting

and satisfy the needs of the painter to practice his art. In a like manner, the soldier cannot practice his art without some technical knowledge. Nor can the supporter of the soldier, the technologist or analyst of war, satisfy the needs of the soldier without knowledge of the technical aspects of war. This book, then, is of interest to both the soldier and the technologist of war.

Most books about warfare are historic in nature, ranging from memoirs such as Xenophon's **The Persian Expedition**[4], Gaius Julius Caesar's **The Conquest of Gaul**[5], and Donn Albert Starry's **Armored Combat in Vietnam**[6], through tactical and strategic treatises such as Frederick the Great's On **the Art of War**[7], Jomini's **The Art of War**[8], and von Clausewitz's ***Vom Kreig*** **(On War)**[9]. (The latter category seldom seen on bookstore shelves.) Some historical analysis of warfare has found its way into print, ranging from Dehlbrück's **History of the Art of War**[10] to Trevor Dupuy's **Numbers, Predictions and War**[11]. The modeling of warfare has its origins in the analysis of history. This is amply evidenced in Lanchester's **Aircraft in Warfare: The Dawn of the Fourth Age**[12], Osipov's articles[13], and Fiske's **The Navy as a Fighting Machine**[14]. (Discussed briefly in Chapter II.) Books on the technical aspects of the modeling of warfare are rare, the exceptions being Dupuy's book and Taylor's **Force-on-Force Attrition Modelling**[15], the former describing an empirical approach from historical data which sadly, despite its aesthetic form, lacks any theoretical foundation which admits the introduction of technological advances (which as Ferrill notes increasingly dominates the nature of warfare,) and the latter giving no attention to historical insights and scant attention to the underlying mechanics of attrition processes.

What this book is, is a combination of historical (both in the classic sense and in the sense of the discipline) and technical (mostly the latter) analyses of warfare models. The approach is somewhat mathematical. A knowledge of the integral calculus and elementary probability theory is assumed; that level of sophistication seems to be the minimum requisite to consider the subject in depth, and is probably enough to dissuade the average professional soldier from reading further. That is not altogether a misplaced view; as I have said, the practice of war is an art form and this book is not primarily concerned with the art form. However, to borrow a model from my own profession, a physicist will gather knowledge (and tools) from a mathematics book without having the depth of understanding of the proofs of the theorems that is required of the mathematician. Rather, the physicist largely accepts the proofs' validity at face value and uses them as tools in the practice of his profession. So too may the soldier make use of the material in this book, using the results of the derivations and analyses as tools. Fundamentally, however, this book is concerned with the analysis of warfare, and as such, is of interest to the more analytically inclined of the community. This book will not likely ever be mentioned in the same breath or be of comparable value to the soldier as Vom Kreig; it may be of some use to the technical

community that supports the soldier in the pursuit of his art form by providing him with the materiel and doctrinal tools of his trade.

Of necessity, much of this book is concerned with the attrition process of warfare. This is partly due to our fascination with that aspect of war, and partly due to the preponderance of the literature on that aspect. Presumably, attrition is more amenable to analysis than other aspects of war!

While this book does not portray itself to be an historical work, some data on historical warfare is included to address salient points in the mathematical theory and provide insight in the analyses. Of necessity, those data are limited; warfare is not, and never will be, a strictly scientific subject. We cannot conduct scientific experiments on warfare. The control problems aside, moral and economic factors preclude such experiments. As a result, considerable uncertainty must and does exist in the historical data. Of necessity then, the data must be culled. (I am not an historian, and detractors may claim that I have been overselective in my choices or have been deficient in the exhaustiveness of my scholarship. I cannot defend myself on the historical selections except to state that I have attempted to be honest in my selections.) In many cases, the culling of historical data is dictated by the requisites of the mathematics - a minimum of numeric data is necessary and only battles for which that minimum can be found can be subjected to analyses of the types presented here. Much of this data, as I have stated, is uncertain; in particular, meaningful data on the actual duration of the vast majority of battles is wanting, or at best, suspect. Even force strengths are uncertain, with contradictory reports often being the norm.

Within these limitations, this book presents few conclusions. Rather, it attempts to lend insight into the dynamics of warfare. The reader should remember that this is an immature discipline. It has few laws and is predictive in only the most cursory sense. (We do not mean here the Laws of War; they are the laws of the art of war, not laws in a technical sense.) Still, the discipline offers considerable promise in terms of developing into something which will be a contributor to man's understanding of his universe. May this book serve in some manner to hasten that day.

There is, I hope, a wider audience for this book than the professional soldier. The core group for which this book is written are the students, those who practice the peripheral professions of war and must learn their trade and continually update and expand their understanding of it. These students include the developers of weapons and doctrine, the analysts and users of combat simulations, the civilian and soldier managers of military programs, their counterparts in the defense industry, and even academia.

Preface

I would also like to acknowledge the encouragement and professionalism of Dr. Griff Callahan (COL USA Ret) of Georgia Tech, and discussions with Jim Taylor, Jim Dunnigan, and COL Trevor DuPuy (USA Ret) who shaped my concepts and understanding of this subject and my resolve to write this book. I would also like to acknowledge the many military practitioners (green suiters) with whom I have been privileged to have discussions which have been educational and provocative.

References

1. Ferrill, Arther, **The Origins of War**, Thames and Hudson, New York, 1985.

2. Gabriel, Richard A., and Karen S. Metz, **From Sumer to Rome**, Greenwood Press, New York, 1991.

3. Jones, Archer, **The Art of War in the Western World**, University of Illinois Press, Urbana, 1987.

4. Xenophon, **The Persian Expedition**, Penguin Books, Inc., Baltimore, 1965.

5. Gaius Julius Caesar, **The Conquest of Gaul**, Penguin Books, Inc., Baltimore, 1984.

6. Starry, Donn Albert, GEN, U.S. Army, Ret., **Armored Combat in Vietnam**, Bobbs-Merrill, New York, 1980.

7. Frederick the Great, **On the Art of War**, The Free Press, New York, 1966.

8. Jomini, Antoine-Henri, Baron de, **The Art of War**, J. B. Lippincott and Company, Philadelphia, 1862, (reprinted by Greenwood Press, Westport, Connecticut, undated.)

9. Clausewitz, Carl von, ***Vom Kreig*** **(On War)**, 1832, Princeton University Press, Princeton, New Jersey, 1976, and Penguin Books, Inc., Baltimore, 1968.

10. Delbrück, Hans, ***Geschichte der Kreigskunst im Rahmen der politischen Geschichte*** **(History of the Art of War in the Framework of Political History)**, Berlin, 1900 and 1962-64, Walter J. Renfroe, Jr., trans., University of Nebraska, Lincoln, 1975 and 1990.

11. Dupuy, Trevor N., COL, U.S. Army, Ret., **Numbers, Prediction and War**, Bobbs-Merrill, New York, 1979.

12. Lanchester, Frederick W., **Aircraft in Warfare: The Dawn of the Fourth Arm**, Constable and Company, LTD., London, 1916.

13. Osipov, M., "The Influence of the Numerical Strength of Engaged Forces on their Casualties", *Voennie Shornik*, 3-7, 1915, translated in the report of the same title, by Drs. Robert L. Helmbold and Allan S. Rehm, U.S. Army Concepts Analysis Agency, Bethesda, MD, Research Paper CAA-RP-91-2, September 1991.

14. Fiske, Bradley A., Rear Admiral, U.S. Navy, **The Navy as a Fighting Machine**, Naval Institute Press, Annapolis, 1988, originally Charles Scribner's Sons, New York, 1916. Includes as Appendix C, Lt. J. V. Chase, "Force-on-Force Effectiveness Model for Battle Lines", Naval War College, 1902.

15. Taylor, James G., **Force-on-Force Attrition Modeling**, Operations Research Society of America, Arlington, 1981, and **Lanchester Models of Warfare**, Operations Research Society of America, 1983.

0. Introduction

0.A Can we Define the Question?

Is there a physics to warfare? Or perhaps a better way of asking that question would be "Is there any scientific basis to describe the processes of war?" That may still not be clear, so we shall examine the two key ideas in the question(s): science and war.

The logical starting point is to examine the dictionary. The Random House Dictionary[1] defines science as:

> "(1) a branch of knowledge or study dealing with a body of acts or truths systematically arranged and showing the operation of general laws."

and war as:

> "(1) a major armed conflict as between nations."

The American Joint Chiefs of Staff defines neither in their Dictionary of Military Terms [2] while our chief competitors (until recently) the Soviet Union, provided a long description definition of war[3] (*voyna*):

> "War is an armed conflict between states (coalitions of states) or between striving antagonistic classes within a state (civil war) to gain their economic and political ends."

Finally, we find that physics[4] is

> "The study of those aspects of nature which can be understood in a fundamental way in terms of elementary principles and laws."

These dry definitions do indeed allow us to ask over questions in a (hopefully) more meaningful way: Are these general laws or elementary principles - operating in the armed conflicts between states? Obviously, I (and others) must have some reason to believe that there are, I would not have written this book advancing to describe some of our knowledge (and offer, our lack of it). Equally obviously, if I do not move on to something a little less dry then their definitions, you, the reader, will cease to read.

0.B Lies, Lies, and Dammed Lies

Our common vision of the scientist is largely shaped by science fiction movies that portray scientists as mad, cackling men (and occasionally women) who perform diabolical experiments without regard for the social consequences of their acts. Part of this is true. Scientists do perform experiments (and develop theory based on these experiments) to uncover and understand the fundamental principles of the universe around us - science and experiments are fundamentally linked! It is not generally true

that scientists are either mad or cackling, and they are generally quite concerned about the social consequence of their efforts - witness the volume of writings and efforts by scientists about the efforts by scientists about the effects and morality of nuclear war. In recent years, considerable consideration has been given to the social and political efforts of science. As Michael Simon[5] notes in his review of Alan Chalmer's book **Science and Its Fabrication** "the distinction between good and bad science is a useful one, but it is not one that can be clearly drawn. The goal of science, as Chalmer understands it, is not certainty but improvement or growth."

War, to most of us and including the soldier, is a terrible thing. That makes its study a paradox. Clearly, we want to study it so that we may avoid its occurrence, or given its occurrence, complete it in as limiting a fashion as possible. That is the universal approach of the modern military professional. The negative side is that if we understand war better, we may apply that understanding to practice it. This paradox is a fundamental example of the two edged nature of knowledge in general and science in particular.

Clearly, war has not been the subject of exhaustive scientific study. There are several reasons for this, and I cannot delude the reader into thinking that my list of reasons is exhaustion. I do believe they are illustrative and reasonably comprehensive, however.

Because of its very terribleness, war does not attract scientists to study it. Nor, are many soldiers scientist or *visa versa*. The nature of the two professions do not allow them to mingle effectively. This does not mean that soldiers do not study war, quite the contrary. Many soldiers are dedicated students and learned scholars of war, but that knowledge tends to be historic and practical in nature. This study, over several centuries, has produced considerable result and theory, but it is a scholarly rather than a scientific type of knowledge. This must not be belittled. This knowledge is important and we shall examine it not only later in this chapter, but throughout this book as well.

As we have already stated, our interest here is the physics of war, or at least of the processes of ground conflict, and this means in particular that we want to examine those processes which are describable in a quantifiable manner. In simple, we want to examine those parts of war that can be described in the exacting vocabulary of mathematics. This is not easily done for two reasons.

First, the soldier is not, as a member of a profession, given to the daily use of math as a tool. Like most of our citizenry, he (or she) is not generally adept at using math as a tool for understanding and describing his world. This results both from our cultural approach to the teaching (or non-teaching - see Appendix X) of math and to practical, accomplishment orientation of his profession. Many professions have this non mathematical character, but that should not preclude the soldier from seeking

greater personal and professional knowledge from efforts like this one.

The second reason, which finally brings us to some meat among all this philosophical rabbit food, is the fundamental linkage between science and experiment. By its very nature, war is not truly amendable to experiment. We cannot, in the interest of science, go into the laboratory and conduct war as an experiment. To coolly conduct measured experiments in war where lives are taken is both ethically and morally impossible. Nor can we make complete use of military field trials and exercises as experiments for two reasons: first, to make detailed measurements of such would completely compromise them - the influence of the observer is disastrous, and second, these trials and exercises are not war and any knowledge that we gain from them is fiercely tainted with uncertainty of the most vicious type. We do not even know what the nature and magnitude of the uncertainties are.

Our only recourse therefore, is history. We can only use what data is available from the battles and wars that have been fought in the past. As we have noted, this is the principle approach of the modern professional soldier - to study the history of war. Can we have however derive scientific knowledge from history?

I will not try to be exhaustive in this introductory chapter, but can sketch none of the most obvious basis for what scientific knowledge we can derive from history and thereby lay a basis for the mathematical theory that we will be describing in the rest of this book.

Until recently, the numerical data on war was not readily accessible, if we can say it is today. There are however, scholarly works of history that describe wars, campaigns, and battles and in these works there are a few numbers. Because of the largely theoretical nature of this book, we have limited our search for historical numbers to sources which compile many battles and looked there for numbers describing the battles. From these compilations, we developed databases of selected battles. The criterion for selecting (and rejecting) battles was very simple - there had to be a minimum amount of recorded numeric data about a battle in the compilation. This culling process is extreme, it reduces the number to something on the order of 1-2 percent of the total battles. Thus, we immediately must view the resulting list with great trepidation, who knows how we have inadvertently slanted and distorted the view that we may derive from these data!

All of these concerns aside, let us at lest look at one, fairly general, set of dates. This one is taken from a historical compilation of battles entitled **Brassey's Battles**,[6] named after the company that published the compilation. We shall describe the source, and the nature of the data base in question detail in Chapter IX, but for now, we are primarily interested in what we may learn about these battles.

This database consists of 107 battles, one of our largest. The earliest battle is

Marathon fought in -490 C.E.. The most recent battle is Goose Green, fought in 1987. The basic data, consisting of the date, combatants, their initial and final (numerical) strengths, and the durations of the battles (in days) are shown in Table 0.1.

Most of these battles were short affairs, lasting a day or less. A few were longer. They represent only a small fraction of the battles fought in the last two-and-a-half millennia, but, as given by our source, they are, for a human and therefore non-exhaustive search, the only ones that have five pieces of numeric information: the initial and final strengths of the two sides, and the duration of the battle. Actually, the culling criteria as somewhat stronger than this - have also culled battles that did not end in a controlled manner - no routs. With one exception, all these battles ended with both armies enact. While there have been many battles that have turned into routs, our intention here is to examine what we may hopefully call normal battles, even if they may not prove with further study to be normal.

Our prescription, for now, will be to examine the contents of this database to see if these are any describable patterns in the data here. We will not attempt, at this time, to perform any type of statistical analysis of these data. What we are interested in are clear patterns that would indicate the possibility of quantitative relationships.

First, examine the way that the force strengths of the battles are distributed by the date of the battle. This is shown in Figure 0.1. The only pattern that we may see here is that most of the battles were fought in the last 500 years. Is this meaningful? Has civilization over the last five centuries become more warlike, or is this the result of better, more thorough recordings. I suspect the latter is our meaningful explanation.

Next, examine the same force strengths, as a function of the battle's duration. This is shown in Figure 0.2. Again, the only pattern that we discern is that most of these battles were short - but we have already noted this. If we look at this data in a log-log plot., Figure 0.3, we see a wider spread of the data, indicating more variation, but no striking pattern. At best, we may only speculate that the shorter battles seem more likely to have the data we need recorded. A leap of speculation could be that most battles only last a day because of the difficulty of fighting at night, and almost all last only a few days because of the intensity of battle.

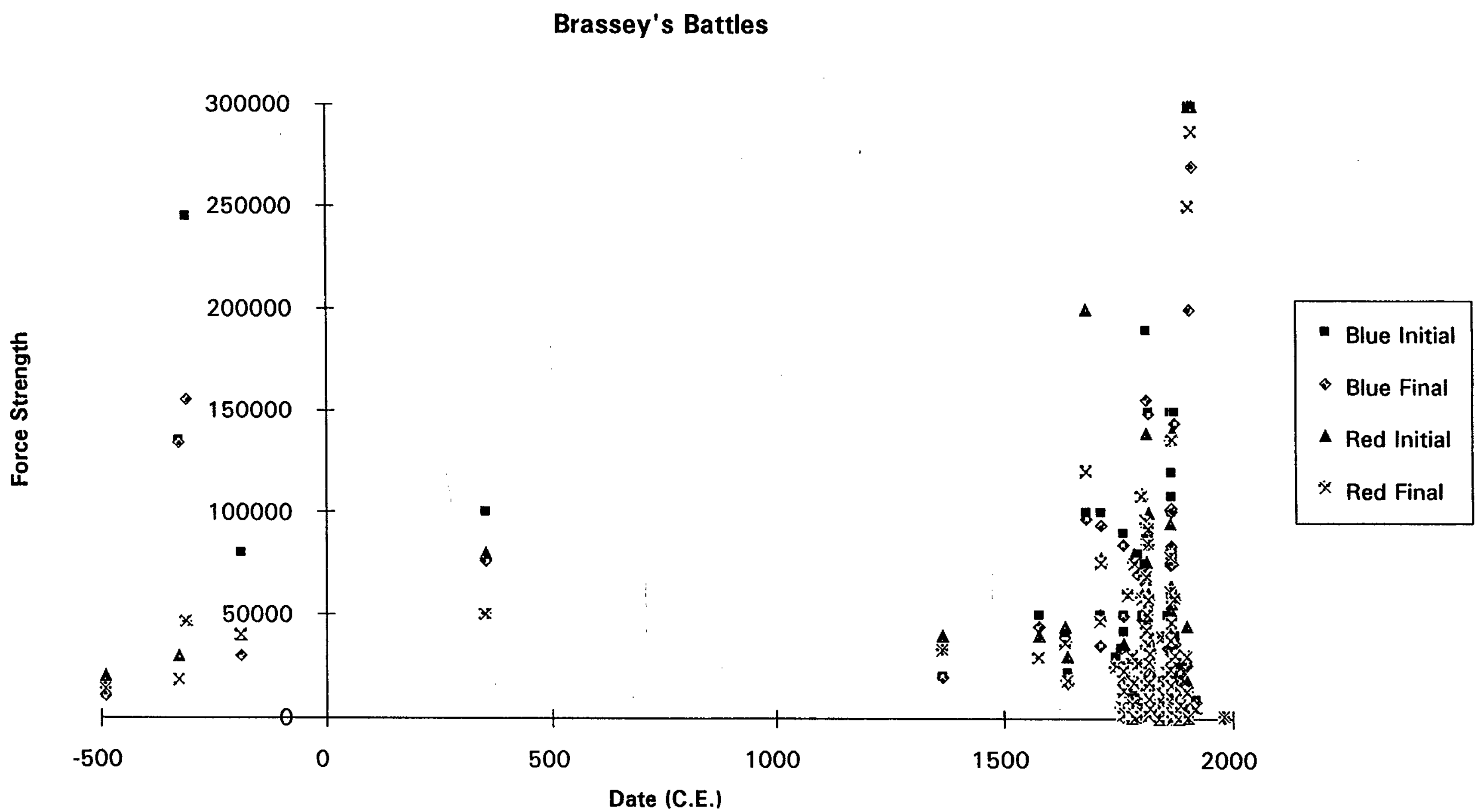
We must recall that even during war, battles are relatively rare. Troops from both sides must concentrate in the same locality at the same time. Except under unusual operational circumstances, the commanders of both sides must want to fight. Since the purpose of these meetings is too often a decision - I win, you loose, - we would expect battles to be intense, and it is this very intensity that will make them rare since the armies cannot fight another battle until they have rebuilt their strength.

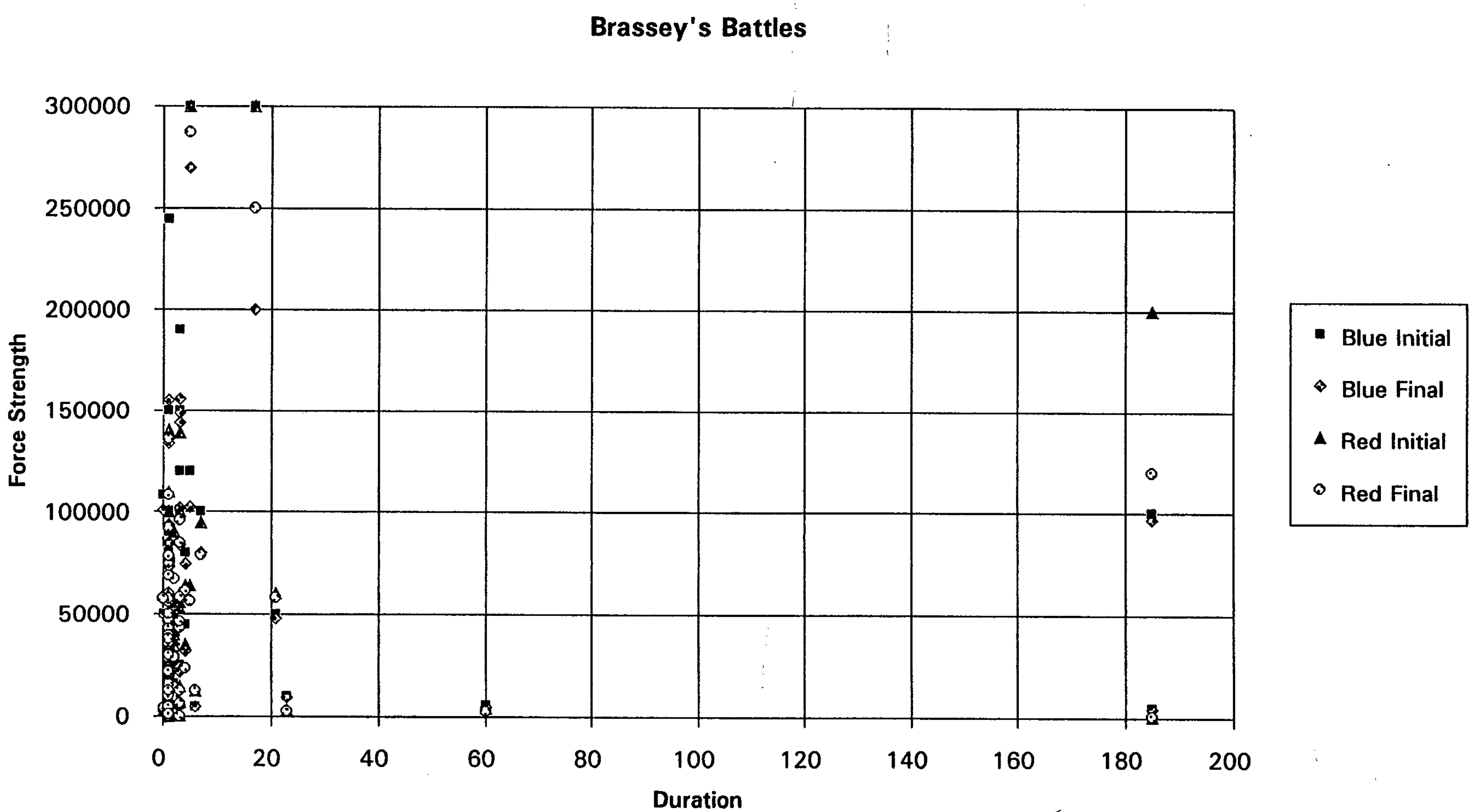
How intense are these battles? To examine this, let us compare the ratio of each sides final to initial strength. This ratio is just the fraction of surviving strength

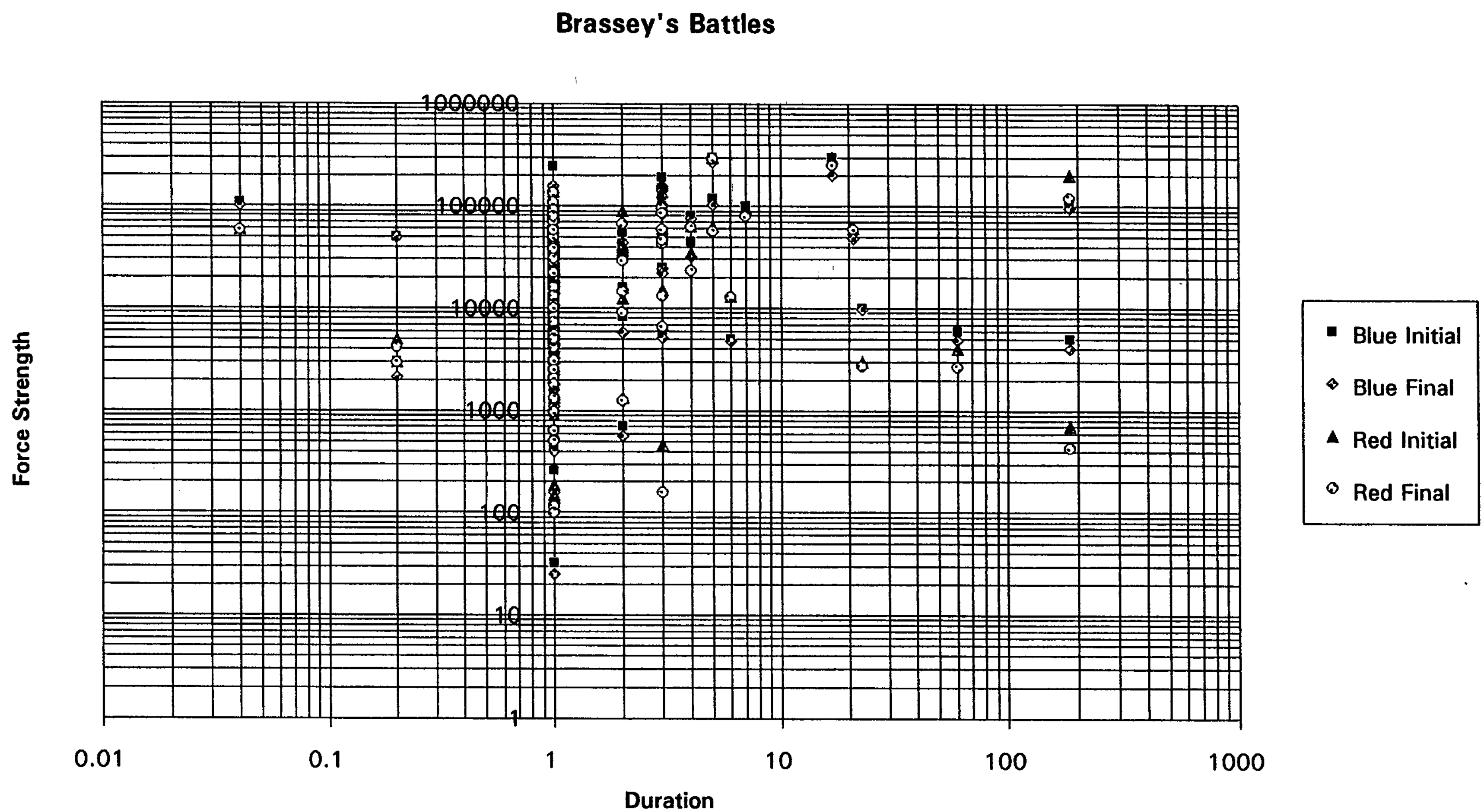
Battle	Date	Blue	BStart	BFinish	BLoss	Red	RStart	RFinish	RLoss	Duration
Marathon	-490	Greeks	11000	10808	192	Persians	20000	13600	6400	1
Hydaspes	-327	Macedonians	135000	134000	1000	Indians	30000	18000	12000	1
Arbela	-311	Persians	245000	155000	90000	Macedonians	47000	46500	500	1
Magnesia	-190	Syrians	80000	30000	50000	Romans	40000	39700	300	1
Mursa	351	Rebels	100000	76000	24000	Romans	80000	50000	30000	1
Navarette	1367	British	20000	19900	100	French/Spanish	40000	33000	7000	1
Marignano	1575	French	50000	44000	6000	Swiss	40000	29200	10800	2
Breitenfeld I	1631	Swedes	40000	37300	2700	Imperials	44000	36000	8000	1
Wittstock	1636	Swedes	22000	17000	5000	Austrians/Spani	30000	18000	12000	1
Zurakow	1676	Poles	100000	97000	3000	Turks	200000	120000	80000	185
Ramilles	1706	French	50000	35000	15000	Allies	50000	47000	3000	1
Oudenarde	1708	French	100000	94000	6000	Allies	78000	75000	3000	1
Mollwitz	1741	Prussians	30000	27500	2500	Austrians	30000	25000	5000	1
Sugar-Loaf Rock	1753	French/Indians	34000	33900	100	British/Indians	6000	5960	40	1
Lake George	1755	French/Indians	1500	1100	400	American Coloni	2500	2098	402	1
Breslau	1757	Austrians	90000	84000	6000	Prusians	25000	20000	5000	1
Plassey	1757	French/Indians	50000	49500	500	British/Indians	3000	2928	72	0.2
Ticonderoga I	1758	French	3600	3223	377	British/America	15000	13056	1944	1
Madras II	1758	French/Indians	6000	4900	1100	British/Indians	4000	2659	1341	60
Zerndorf	1758	Russians	42000	21000	21000	Prussians	36500	22700	13800	1
Plains of Abrah	1759	British	4000	3336	664	French	4000	2500	1500	1
Trincomalee II	1767	British/Indians	12000	8000	4000	Mysore Forces	60000	59840	160	1
Quebec II	1775	British	1800	1782	18	Americans	600	500	100	1
Brooklyn	1776	Americans	11000	9000	2000	British	30000	29690	310	1
Moore's Creek B	1776	Americans	1100	1098	2	British	1800	1770	30	1
Brandywine	1777	British	18000	17410	590	Americans	8000	7100	900	1
Saratoga	1777	British	6000	5400	600	Americans	7200	7000	200	1
Newport	1778	Americans	10000	9689	311	British	3000	2740	260	23
Stony Point	1779	British	700	567	133	Americans	1350	1255	95	2
Panani	1780	British	2500	2413	87	Mysore Forces	18300	17200	1100	1
Hobkirk's Hill	1781	Americans	1551	1281	270	British	900	642	258	1
Sholinghar	1781	British	10000	9900	100	Mysore Forces	80000	75000	5000	1
Martinesi	1789	Turks	80000	70000	10000	Austrians/Russi	27000	26383	617	1
Savannah II	1799	Americans/Frenc	5050	4212	838	British	3200	3045	155	1
Engen	1800	French	75000	73000	2000	Austrians	110000	108000	2000	1
Stekach II	1800	French	50000	48000	2000	Austrians	60000	58000	2000	21
Pultusk II	1806	Russians	37000	34000	3000	French	20000	15800	4200	1
Rio Seco	1808	French	14000	13630	370	Spanish	26000	20000	6000	1

Talavera	1809	British/Spanish	40000	34600	5400	French	50000	42700	7300	1
Aspern-Essling	1809	French	55000	35000	20000	Austrians	90000	67000	23000	2
Eckmuhl	1809	French	90000	85000	5000	Austrians	76000	69000	7000	1
Raab	1809	French	44000	41500	2500	Austrians	40000	37000	3000	1
Wagram	1809	French	190000	156000	34000	Austrians	139000	96000	43000	3
Medelin	1809	French	17500	16500	1000	Spansih	30000	12000	18000	1
Busaco	1810	British	25000	23500	1500	French	40000	35500	4500	1
Albuhera	1811	French	33000	25000	8000	Allies	7000	1800	5200	1
Mohilev	1812	French	28000	27000	1000	Russians	60000	56000	4000	1
Smolensk II	1812	French	50000	41000	9000	Russians	60000	50000	10000	1
Malo Jaroslavet	1812	Russians	24000	18000	6000	French	15000	10000	5000	1
Maya	1813	British	6000	4600	1400	French	26000	24500	1500	1
Sauroren·	1813	French	25000	22000	3000	British	12000	9400	2600	1
Bautzen	1813	French	150000	148700	1300	Prussians/Russi	100000	85000	15000	3
Baltimore	1814	British	3270	2924	346	Americans	17000	16890	110	1
Lundy's Lane	1814	British	3000	2150	850	Americans	5000	4150	850	0.2
Toulose II	1814	British/Spanish	25000	20400	4600	French	30000	27000	3000	1
La Rotheria	1814	French	32000	27000	5000	Allies	100000	92000	8000	1
Arcis sur Aube	1814	French	23000	21300	1700	Austrians	60000	57500	2500	1
Craonne	1814	Prussians	90000	85000	5000	French	37000	31600	5400	1
Quatre Bras	1815	French	25000	20700	4300	British/Dutch	36000	31300	4700	1
Mahidput	1817	Indians	35000	32000	3000	British	5500	4722	778	1
Alamo	1836	Mexicans	2500	900	1600	Texans	185	0	185	1
Maharajapore	1843	Indians	18000	15000	3000	British	40000	39213	787	1
Resaca de la Pa	1846	AMericans	1700	1578	122	Mexicans	5100	4483	617	1
Buena Vista	1847	Americans	4500	3754	746	Mexicans	18000	16500	1500	1
Vera Cruz	1847	Mexicans	5000	4820	180	Americans	13000	12918	82	6
Cerro-Gordo	1847	United States	8500	8100	400	Mexicans	12000	11300	700	1
Alma River	1854	Russians	40000	34100	5900	British/French	26000	22000	4000	1
Inkerman	1854	Russians	50000	38000	12000	British/French	8000	4500	3500	1
Bull Run I	1861	Union	40000	38508	1492	Confederates	30000	28018	1982	1
Wilson's Creek	1861	Union	5600	4364	1236	Confederates	11800	10705	1095	1
Richmond I	1862	Americans	8000	6954	1046	Confederates	6000	5540	460	1
Pea Ridge	1862	Confederates	16000	15200	800	Union	16000	14616	1384	2
Prarie Grove	1862	Confederates	11000	10000	1000	Union	10000	8700	1300	1
Seven Days	1862	Confederates	100000	80000	20000	Union	95000	79000	16000	7
Shiloh	1862	Confederates	43000	32303	10697	Union	42000	28913	13087	2
Fort Donelson	1862	Union	25000	22168	2832	Confederates	12000	10000	2000	1
Fredericksberg	1862	Union	150000	136229	13771	Confederates	80000	78200	1800	1

Front Royal	1862	Union	1063	159	904	Confederates	16000	15950	50	1
Secessionville	1862	Union	6000	5400	600	Confederates	2000	1800	200	1
Nashville	1863	Confederates	31000	29500	1500	Union	41000	37939	3061	1
Chancellorsvill	1863	Union	120000	102000	18000	Confederates	53000	43000	10000	3
Chattanooga	1863	Union	80000	74527	5473	Confederates	64000	61479	2521	4
Stones River	1863	Union	45000	32094	12906	Confederates	35000	23261	11739	4
Brices Cross Ro	1864	Confederates	3500	3008	492	Union	8000	7283	717	1
Monocacy River	1864	Confederates	14000	13300	700	Union	6000	4120	1880	1
Sabine Cross Ro	1864	Confederates	8300	5800	2500	Union	12500	9000	3500	2
Cold Harbor	1864	Union	108000	101000	7000	Confederates	59000	57500	1500	0.04
Mobile Bay	1864	Union	5500	5181	319	Confederates	470	158	312	3
Spotsylvania	1864	Union	101000	83600	17400	Confederates	56000	46400	9600	3
Wilderness	1864	Union	120000	102334	17666	Confederates	64000	56250	7750	5
Winchester III	1864	Union	38000	33117	4883	Confederates	12000	9897	2103	1
Custozza II	1866	Austrians	80000	75400	4600	Italians	140000	136168	3832	1
Mantana	1867	Italians	10000	8900	1100	Allies	5000	4818	182	1
Fort Kearney	1867	United States	32	25	7	Indians	1500	1300	200	1
Bel fort II	1871	French	150000	144000	6000	Germans	60000	58000	2000	3
Saint Quentin I	1871	French	40000	36500	3500	Germans	33000	30600	2400	1
Ulund	1879	Zulu	20000	18500	1500	British	5000	4907	93	1
Rorke's Drift	1879	Zulus	4000	3600	400	British	139	114	25	1
Bronkhurst Spru	1880	British	259	104	155	Americans	150	98	52	1
Son Tai	1883	Chinese	25000	24000	1000	French	7000	6590	410	3
Slivnitza	1885	Serbs	25000	22000	3000	Bulgars	15000	13000	2000	3
Omdurhman	1898	British	26000	25500	500	Mahdists	45000	30000	15000	1
Atbara	1898	British/Egyptia	14000	13430	570	Mahdists	18000	13000	5000	1
Mafeking	1900	Boers	5000	4000	1000	British	700	427	273	185
Mukden I	1905	Russians	300000	200000	100000	Japanese	300000	250000	50000	17
Tannenburg II	1914	Russians	300000	270000	30000	Germans	300000	287000	13000	5
Maypo	1918	Chileans	9000	8000	1000	Spanish	6000	5000	1000	1
Goose Green	1982	British	450	402	48	Argentines	1350	979	371	1


Brassey's Battles
Force Strength
300000
250000
200000
150000
100000
50000
0
-500
0
500
1000
1500
2000
Date (C.E.)
Blue Initial
Blue Final
Red Initial
Red Final


Brassey's Battles
Force Strength
300000
250000
200000
150000
100000
50000
0
0
20
40
60
80
100
120
140
160
180
200
Duration
Blue Initial
Blue Final
Red Initial
Red Final


Brassey's Battles
Force Strength
1000000
100000
10000
1000
100
10
1
0.01
0.1
1
10
100
1000
Duration
Blue Initial
Blue Final
Red Initial
Red Final

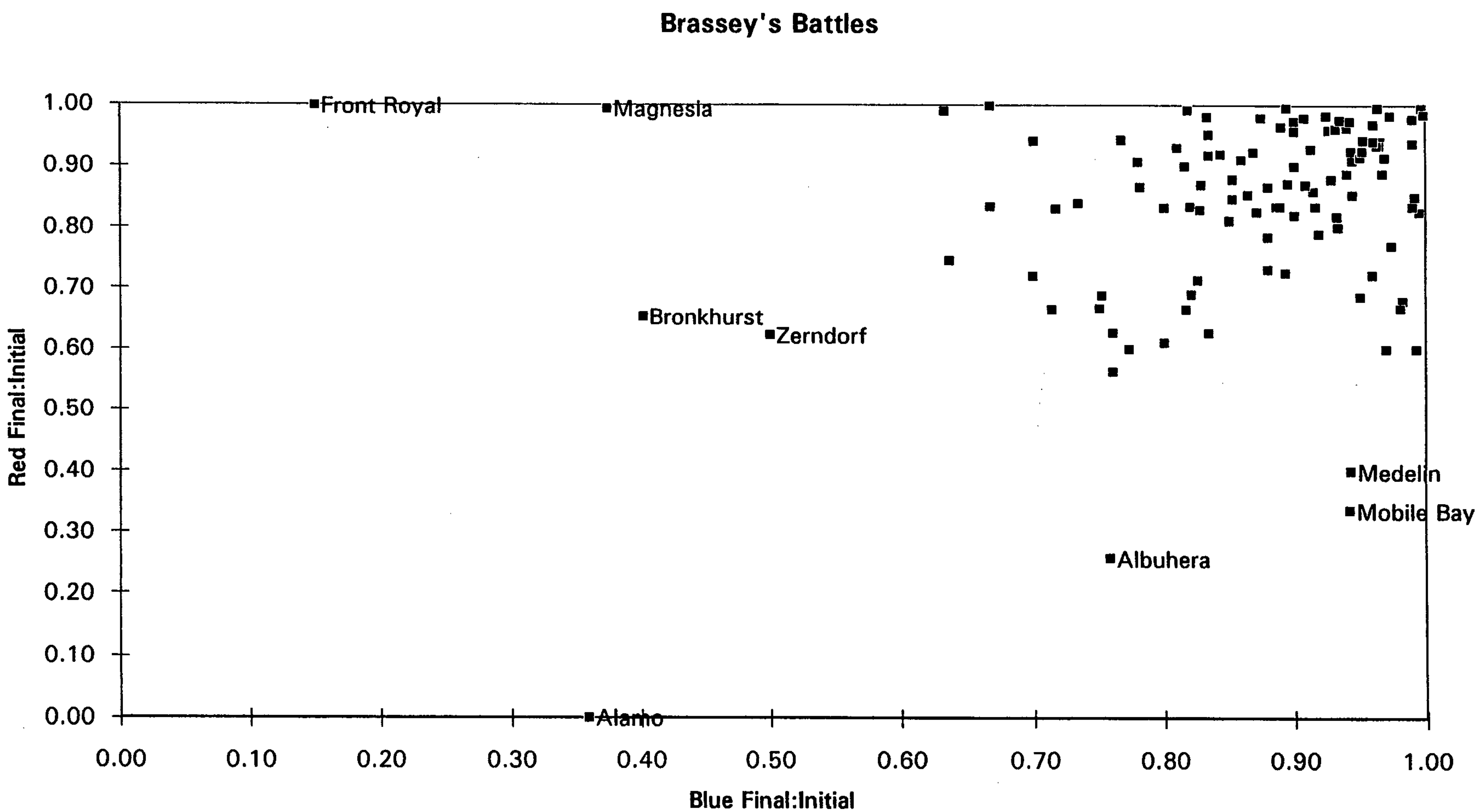
at the end of the battle. The forces in each battle have been arbitrarily divided into two sides - Red and Blue. (Actually the division was decided by the authors of the historical collection. The first side named in the description of the battle is the Blue side.) I have made no attempt to determine which side won or lost, merely divided them.

The cross plot of these ratio's is shown in Figure 0.4. Note that except for the eight battles explicitly called out, none of thee battles end with either force having a surviving force less than 60 percent of its original strength and the majority were considerably more. Clearly then, only 8 out of 107 (less than 10 percent) of these battles could be said to be particularly vicious.

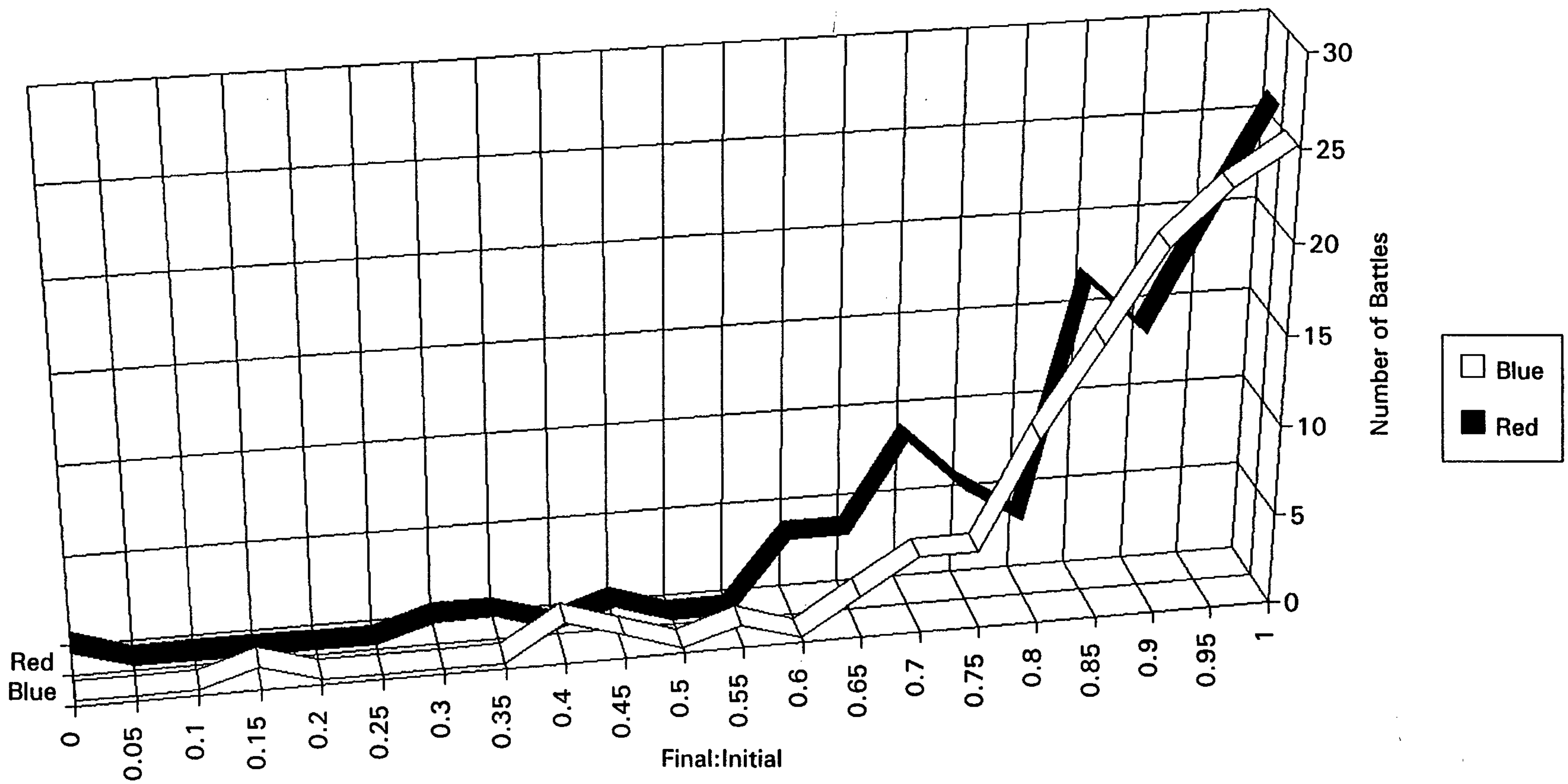
But what is vicious? At the personal level, the two most obvious question are "Did we win?", and "Was I/my friends/relations killed or wounded?" The first question we will consider later in the book when we examine theories of winning. The second question is one that we must inure ourselves against. As callous as it may seem, our approach here must be to accept that some of the troops engaged in a battle will be killed and content ourselves at this time with how many, functionally, that are not. If we do just that, and take the data plotted in Figure 0.4, and ask in how many battles was the servicing fraction between x and x + →x, we get the distribution in Figure 0.5. There are two distributions, one each for the blue and red forces. These curves show what we had surmised from Figure 0.4, that few of these battles took more than a 40 percent toll in strength.

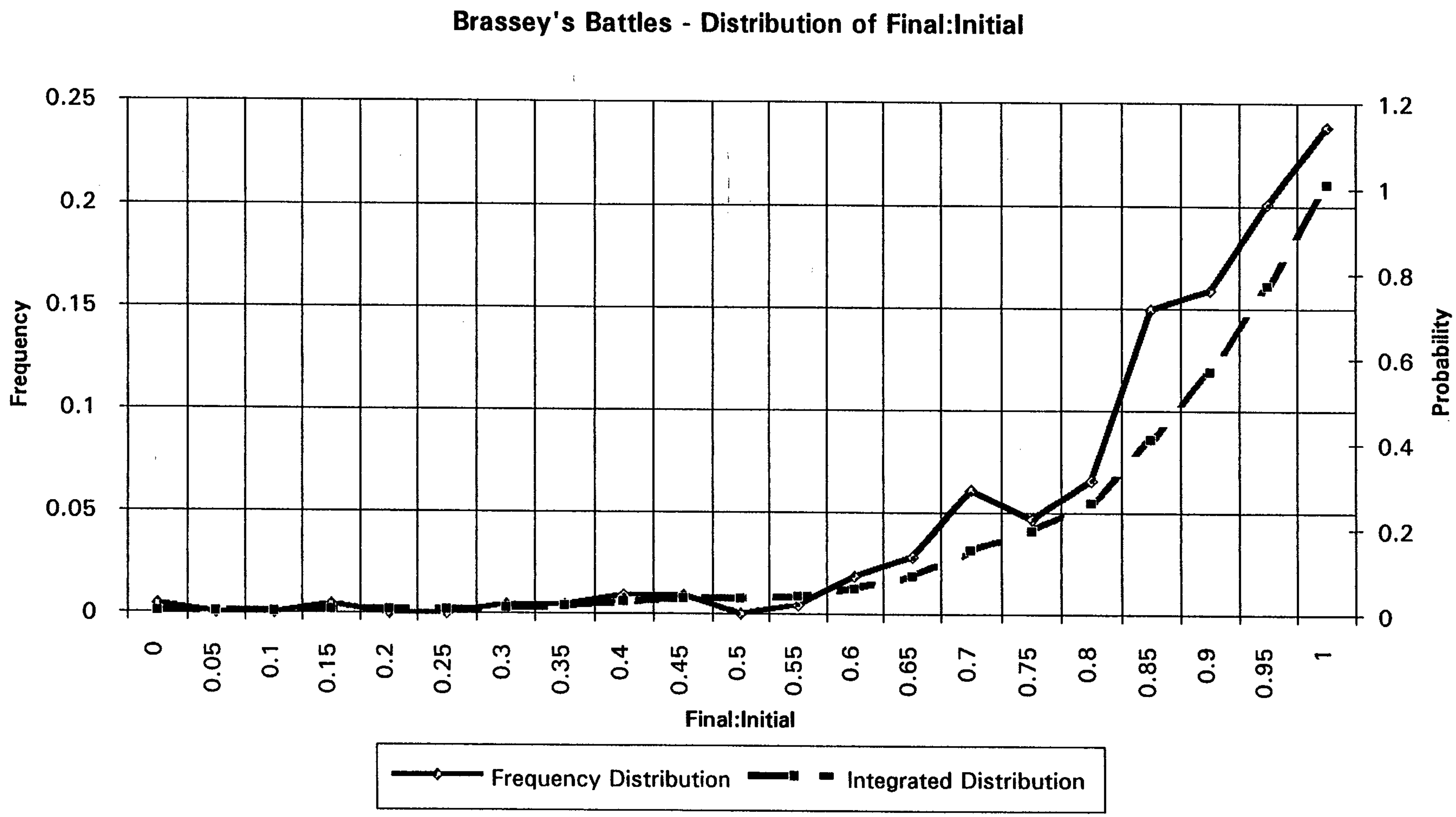
We note that these two distributions are not identical, but are similar in shape. Since the assignment of which force was on each side was arbitrary, we should not expect any strong relationship between them. But if these two curves are similar, may they not be perceived as two sets of random samples from the same distribution? At this point, we have no reason not to view them as such, and to combine the two curves. If we do this, and divide by the total number of forces (twice the number of battles) we get the solid curve in Figure 0.6, which is just the joint frequency distribution of the surviving fraction for these 107 battles.

The dashed curve in this figure is the integral of the frequency distribution. This curve is obviously a negative exponential of the loss fraction (= 1- surviving fraction.) We may read the curve in the following manner: for any given surviving fraction value, the probability that corresponds to the curve is the probability that the surviving fraction will be smaller than the surviving fraction value. For example, there is a 20 percent probability that the surviving fraction after a battle (if we accept this data as representative) will be 75 percent or less. Similarly, there is a 50 percent probability that the surviving fraction after a battle will be about 87.5 percent or less (and obviously, an equal probability that it will be more.). While this is surely a lot, it is a great difference from the view of battles as duels to the death. Clearly, the preservation of the force, if not of individual life, is a major consideration in these


Brassey's Battles
Front Royal
Magnesia
Bronkhurst
Zerndorf
Medelin
Mobile Bay
Albuhera
Alamo
Red Final:Initial
Blue Final:Initial
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

Brassey's Battles - Frequency Distribution




Brassey's Battles - Distribution of Final:Initial
Frequency
0.25
0.2
0.15
0.1
0.05
0
Probability
1.2
1
0.8
0.6
0.4
0.2
0
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1
Final:Initial
Frequency Distribution
Integrated Distribution

normal battles.

Let us now turn to examination of the structure of the loses and final force strengths in terms of the initial force strengths. These data are plotted in Figure 0.6 for the Blue force. While there is no striking pattern for the loses, then most certainly is for the final force strengths. As we may see, for initial force strengths, less the about 150,000, there is a clear upper edge to the data, and very little spread to the data as a whole. This behavior is repeated for the Red force, shown in Figure 0.7.

To confirm this, we again combined the two sets of data. In Figure 0.8, we plot the loses to both sides in all battles, and find no obvious pattern. In Figure 0.9; however, where we plot the final force strengths, there is an obvious pattern. Clearly, there is an upper edge which looks remarkably straight, which seems to set an upper limit on how much the final force strength is, and a less obvious, but still strong indication of a lower limit. Further, there is very little spread to the data. On the basis of just the knowledge that the final strength must be no greater than the initial strength, we would expect the lower triangular half of the graph to be peppered with data points. The skeptic may be tempted to advance that the sharp upper boundary is just the straight line across the graph, but closer examination will show this apparent straight line to have a slope less than one.

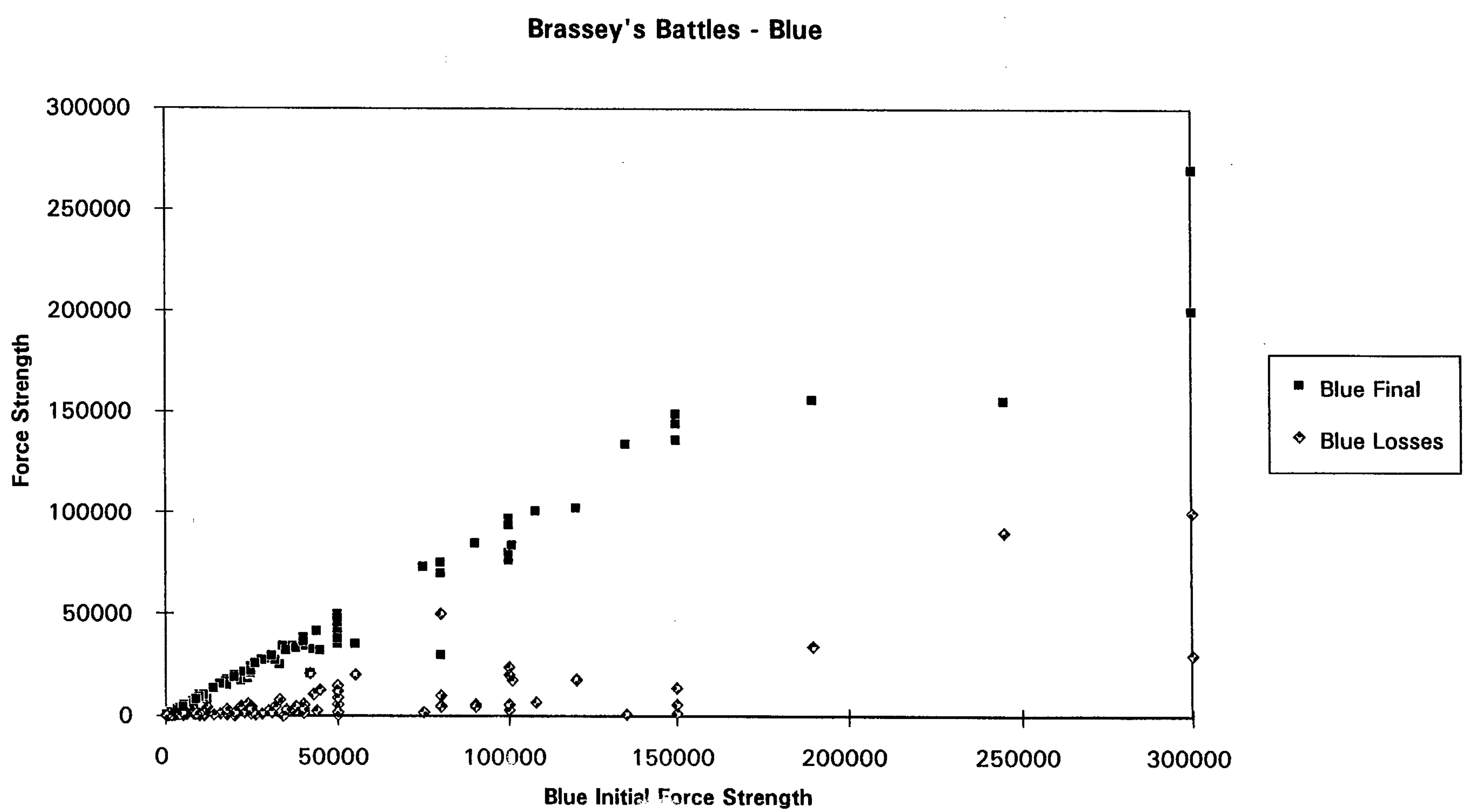
This data represents an historical foundation for a physics of warfare. Clearly, the final force strength must be viewed as being functionally dependent on the initial force strength. Equally clearly, it must also depend functionally on other factors, but it is not obvious what these are from the figures we have presented here.

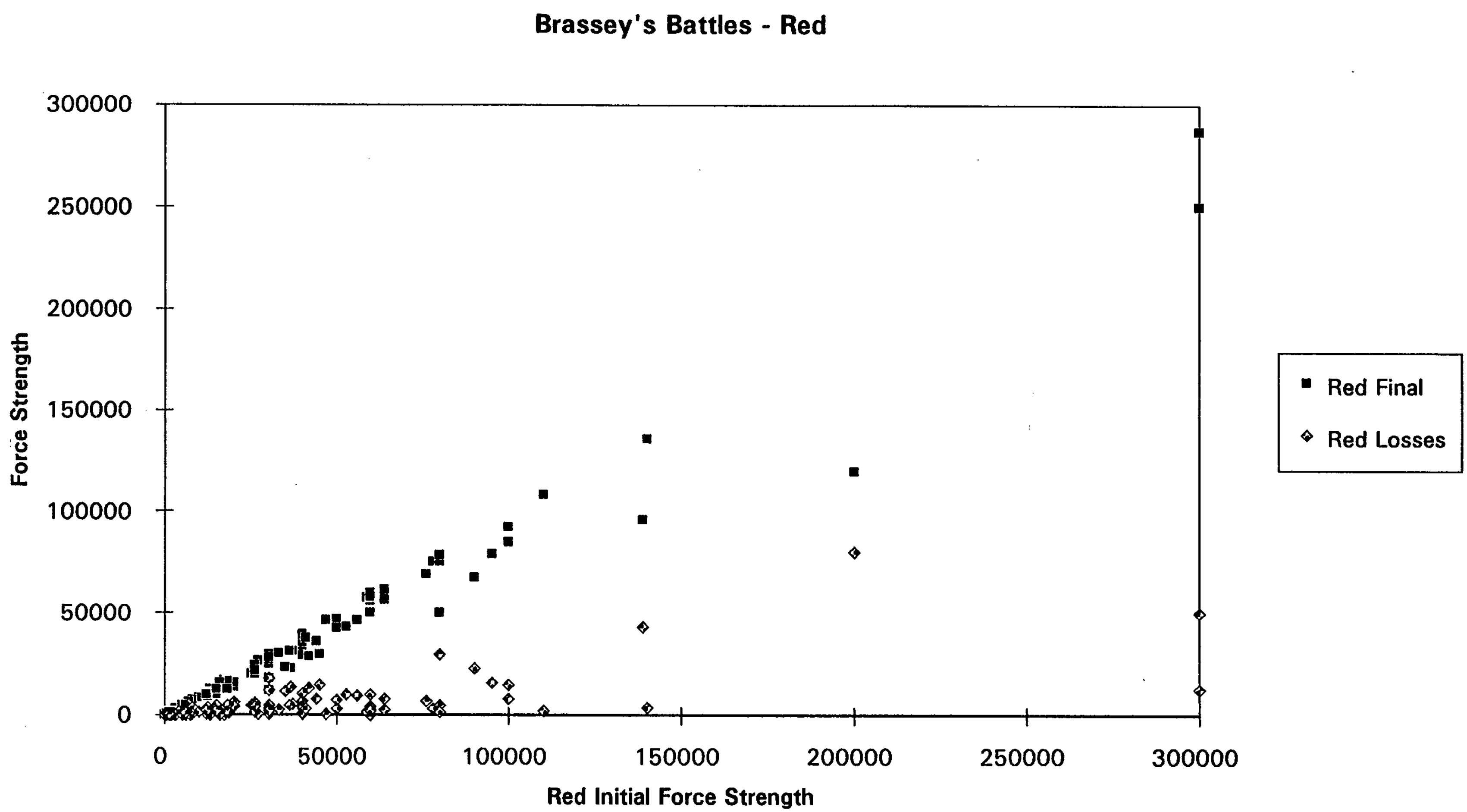
In the following chapters, we shall develop the mathematical basis of one of the most compact of the few theories of attrition, that of Lanchester. Having done this to a reasonable level, we will then reexamine these and other historical data in light of the theory of attrition. Having made this comparison, we shall then broaden our scope to examine other theories of attrition and warfare, and examine the uses of the theory in practice.

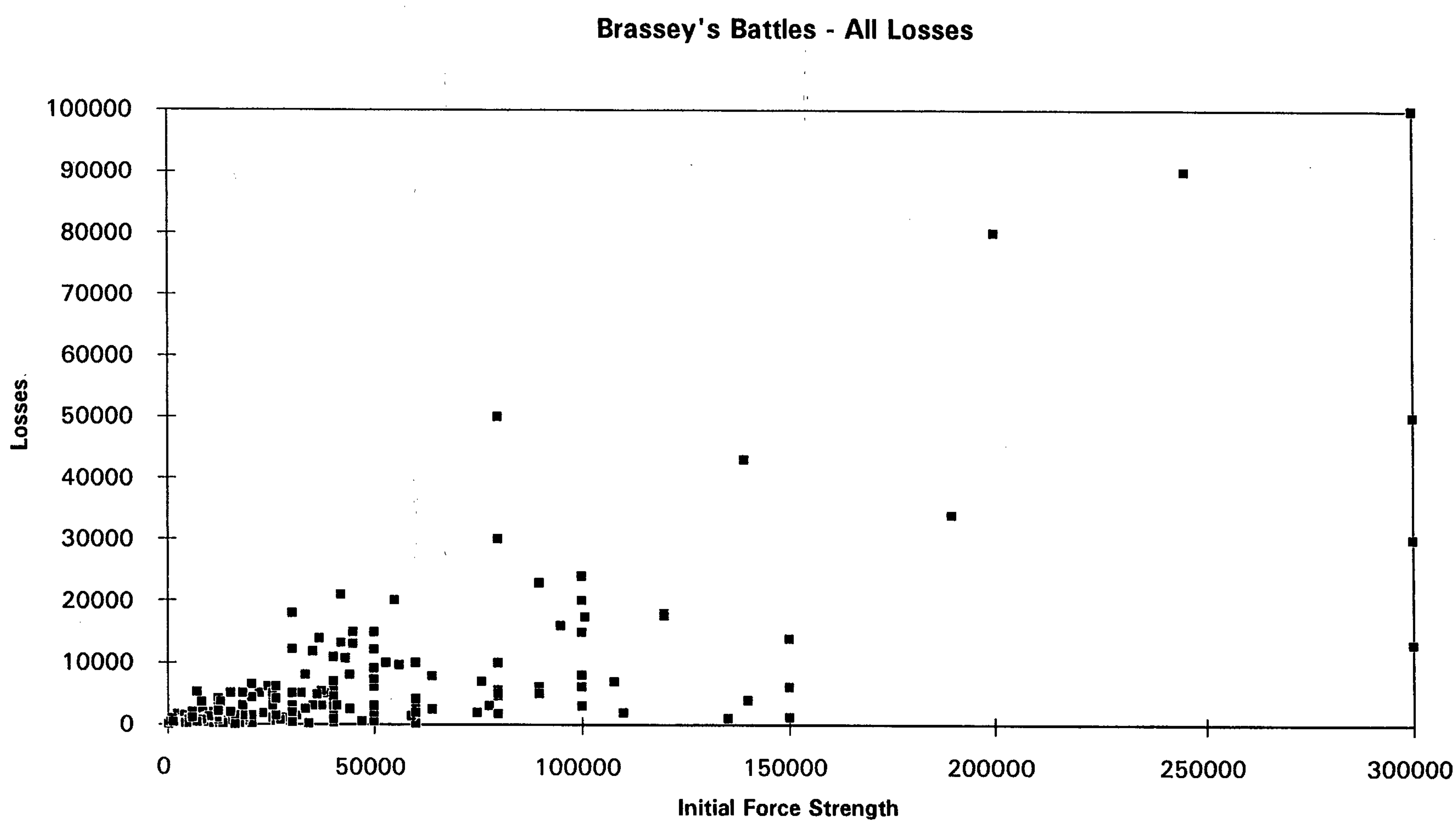
Before commencing on this mathematical journey of theory, we shall finish this chapter with a brief discussion of some of the fundamental principles of war as developed by centuries of scholarly study.

0.C The Principles of War

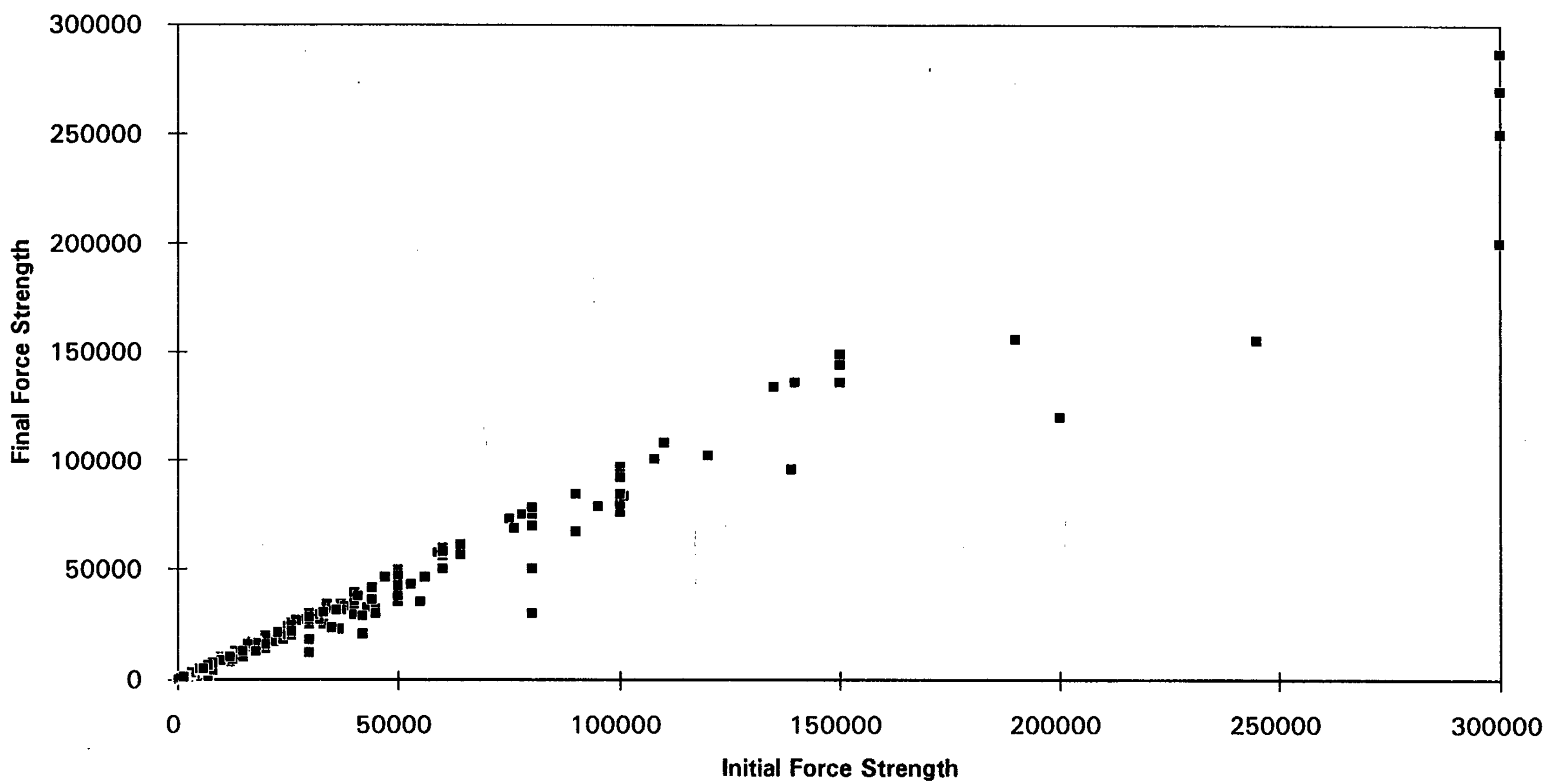
The historical study of war by soldiers and historians has taken the form of many theories of tactics, strategy, and rules of war. Previous attempts have even tried to associate the use of mathematical models in understanding warfare. These attempts, notably those of Jomini, have been roundly denounced by even more students, notably Clausewitz himself. This debate appears to have at its heart the


Brassey's Battles - Blue
Force Strength
300000
250000
200000
150000
100000
50000
0
0
50000
100000
150000
200000
250000
300000
Blue Initial Force Strength
Blue Final
Blue Losses


Brassey's Battles - Red
300000
250000
200000
150000
100000
50000
0
Force Strength
0
50000
100000
150000
200000
250000
300000
Red Initial Force Strength
Red Final
Red Losses


Brassey's Battles - All Losses
Losses
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
0
50000
100000
150000
200000
250000
300000
Initial Force Strength

Brassey's Battles - All Finals



fundamental values of the soldier. The adherents of these theories and methods see them as useful for understanding the practice of war. By their nature they attempt to reduce the environment of the battlefield and the theater of war to simple chunks that can be analyzed. This simplicity is the root of their critics' complaints that these chunks are too simple, are unrealistic, and misleading. This illustrates the fundamental difference between the practical necessity imposed on the field soldier and the ivory tower, start simple and improve, approach of the scientist.

Interestingly enough, despite the aversion to the application of quantitative analysis to war, all students of war advance some form of analytical discipline. In Clausewitz's case, it is called Critical Analysis or *Kritik*. We must conclude that the soldier is not blind to the value of analysis, but will always temper his valuation of it to its accuracy and applicability in his sternly pragmatic world view. The ultimate test of the scientist applies strongly here - is it accurate? In this regard, there is common ground.

Of the analyses conducted over the years, the most profound products have been the Principles of War. These principles are the direct result of the evolution of decisive, persisting (in Jones' terminology) and even total war that has evolved on the past few centuries. Their applicability beyond the scope of conventional warfare to non conventional, guerrilla, or even economic warfare has been argued, not without elements of general validity. Even Clausewitz, who decried the tendency to view warfare in terms of fixed rules because of his vision of its chaotic nature, found some guiding principles to be necessary for any comprehensive theory. Because of their fundamental importance to forming a vision of warfare, we present them here in a modified form as they appear in the U.S. Army's Field Manual 100-5, **Operations:**[7]

> The Principle of the **Objective** states that every military operation be directed towards a clearly defined, decisive, and attainable objective. This means that no action should be taken in warfare that does not have some definite, even explicit, goal, and that that goal be meaningful, and attainable. Obviously, this implies that there is some common plan for the force and that that plan is shared, clear, and realistic.
>
> The Principle of **Offensive** states that the most effective and decisive way to attain a goal is by taking offensive action and/or by maintaining the initiative. Initiative is a concept based on the idea of being able to take actions that force one's opponent to respond to, rather than the other way about., Offensive action, which may be strategic and/or tactical, is viewed as being decisive. While it has been possible to maintain initiative while being purely defensive, these cases are viewed as being historically rare. Of note is the special case of being strategically offensive while being tactically defensive. The situation leading to the Battle of the Alamo is an example. This principle obviously builds on the Principle of the Objective.

The Principle of **Mass** (or Concentration) states that combat power should be concentrated at the decisive place and time. In this case, combat power may not just mean superior numbers, but superior fighting capability. This principle does not suggest that forces should be concentrated all of the time. It does suggest that forces should be concentrated at the right place at the right time to achieve decisive results. This principle is most obviously linked to the theories of attrition that we shall describe in subsequent chapters.

The Principle of **Economy of Force** states that only minimal combat power should be allocated to secondary efforts. This means that the army may be divided to pursue several goals but that the Principle of Mass should apply for the primary effort. While clearly permitting division, this principle surfaces the difficulty of knowing exactly what effort will be primary while providing enough force to achieve the secondary goals.

The Principle of **Maneuver** states that one's enemy may be placed in a position of disadvantage through the flexible application of combat power. At a superficial level, this principle seems to suggest that by moving one's forces, the enemy is placed at a disadvantage, thus maintaining or seizing the initiative. This principle means this, of course, but it also implies a flexibility to move and realign one's forces.

The Principle of **Unity of Command** states that there be only one responsible commander who direct the efforts to achieve an objective. This Principle addresses a question of biblical importance, how to serve two masters? Additionally, it applies the Principles of Mass and Economy of Force, suggesting the necessity of common goals, clear objectives, and a rigorous chain of responsibility.

The Principle of **Security** dictates that one must not allow the enemy to acquire an unexpected advantage. In one sense, this is the opposite of the Principle of Maneuver applied to one's own forces - don't allow the enemy to gain advantage. At the same time, it states that initiative must not be lost, and pragmatically, Don't Be Surprised!

The Principle of **Surprise** states that it is desirable to strike one's enemy at a time and/or place, and/or in a manner that he is physically or psychologically unprepared for. This is the reciprocal of the Principle of Security.

The Principle of **Simplicity** states that plans should be clear and uncomplicated and that orders be concise and understandable. This is a pragmatic reinforcement to the Principle of the Objective. It is dignified with a special acronym: KISS - Keep It Simple, Stupid!

These are the Principles of War. Other nations have other sets of principles, but these tend to have great commonality in form and content, if not number and name. These principle form a fairly comprehensive set of rules for conducting military operations, although we see that they might equally well be applied to many human activities. They are a set in the sense that they are interrelated and reinforcing. They are not quantifiable, and are analytic only in a subjective sense. Because they are an embodiment of a theory of warfare however, they must be considered in any quantitative formulation of war, and we shall refer to them during the progress of this book.

References

1. Flexner, Stuart Berg, editor in chief, **The Random House Dictionary, Random House**, New York, 1980.

2. Joint Chiefs of Staff, **U.S. Department of Defense Dictionary of Military Terms**, Greenhill Books, New York, 1987.

3. Radziyevskiy, Professor Gen-Col A.I., sr. ed., **Dictionary of Basic Military Terms**, 1965, DGIS Miltilegical Section Translation Bureau, Secretary of State Department, Ottawa, Canada, trans., published by the United States Air Force as VOL. 9 of Soviet Military Thought, U.S. Government Printing Office, Washington.

4. Lapedes, Daniel N., **McGraw - Hill Dictionary of Physics and Mathematics**, McGraw-Hill Book Company, New York, 1978.

5. Simon, Michael, American Scientist, VOL 80, p 404.

6. Lafflin, John, **Brassey's Battles**, Brassey's Defense Publishers, London, 1986.

7. Headquarters, Department of the Army, **Operations**, Field Manual Number 100-5, Washington, D.C. 5 May 1986.

I. Definitions and Background

I.A. Introduction

This book is not *per se* a text on modeling in the general sense. This chapter however, is included to provide the reader with either a (bare) minimum of background on the general topic, or a commonalty of understanding of definitions, descriptions, etc.

My terms do not, in general, always agree with those in common use in different sections of the community. The techniques and vocabulary of modeling (and simulation and gaming) in physics are different from those used in other disciplines (such as Operations Research.) Even those who are familiar with the discipline will quickly find that I have not been loath to invent new terms or develop new techniques when the occasion warrants and I could find no historical usage.

I.B Definitions and Descriptions

The description of the mechanics of warfare is a task in modeling and simulation. These latter terms are frequently used synonymously; the purpose of this section is to address the definitions and descriptions of these terms.

First, a **model** is a mapping of reality into comprehensibility. A model may always be expressed in informational symbology, which include, but are not limited to, words and mathematics. Griff Callahan of Georgia Tech uses the definition:

> "Modeling is creating representations of specific human perceptions of reality, using imitative or analogous physical or abstract systems to serve as a basis for language."[1]

In simplest terms, a model is a representation of some aspect of reality in terms which can be absorbed and manipulated by the human mind.

In general, a model will only express one facet of reality, although that facet may be complex. Ideally, a model will also be invertible or reversible. Unfortunately, this is not always the case.

The process of developing a model is known as modeling.

A **simulation**, on the other hand, is a tool for expressing the world in understandable terms, constructed from one or more models and a set of logical rules for relating models' interaction. Dr. Callahan uses the definition:

> "Simulation is the use of computers or other devices as tools for experimentation with models."

Finally, **gaming** is the use of one or more simulations to gain understanding or insight. Dr. Callahan uses the definition:

> "Gaming is simulation involving human operators in a competition played according to rules and decided by superior skills or good fortune."

To illustrate the differences among these three, let us consider the use of a 'model' airplane and a wind tunnel to understand the flight characteristics of the 'real' airplane. The 'model' airplane is a model of the real airplane, and the wind tunnel is a model of the environment that the 'real' airplane operates in - the atmosphere. The combination of 'model' airplane and wind tunnel are a simulation of the flight of the 'real' airplane in its 'real' environment. The use of the simulation - operation of the wind tunnel with the 'model' airplane in it - is gaming of the flight of the airplane.

In summary, we may distinguish among models, simulations, and gaming by their nature. A model is abstract, a simulation is concrete (in the sense of an implementation of one or more models,) and gaming is active (the simulation (or tool) is used.) Although we may, and many practitioners do, use these terms almost synonymously, we shall attempt to make distinctions among them. Models are representations of reality while simulations are collections of one or more models for experimental or calculational purposes. Fundamentally, simulations are used to generate numbers from the models.

In practice, the distinction among the three terms becomes indistinct. What does remain distinct however, are the actions associated with these:

- modeling is the development of models,
- simulating is the construction of simulations, and
- gaming is the pursuit of understanding.

{Being the author allows me from time to time to insert extraneous and even outre comments, often on nonquantifiable subjects such as morals or ethics. Many of these comments are my opinions, but being in charge, they appear in black and white, and the unwary reader may erroneously decide that I am passing on arcane knowledge. Sometimes this will be the case; other times, I will only be relating war stories or expressing sour grapes.

One of the loudest of my pet peeves is the question of documentation of a model versus documentation of a simulation. The documentation of a model should be complete enough that a simulation can be constructed embodying it. The documentation of a simulation should be sufficiently complete that the documentation of the model(s) can either be found or is included, and the logical interplay of the model(s) in the simulation is fully explained. Alternately, the documentation of a simulation should be sufficiently complete that the average simulationist can reconstruct the simulation from the documentation (and its references.) To my mind

this is the ultimate test of documentation; completeness for a simulation - can someone who has never used the simulation, or built a similar simulation, build this simulation? If the answer is not yes, then the documentation is inadequate.}

I.C. A Distinctive, Illustrative Example

A model of a gunman's performance may be a probability of kill as a mathematical function of several variables such as the accuracy of the aim and fire, the muzzle velocity, the shape and mass of the bullet, the range to the target, the atmospheric density and wind velocity, the intensity and spectrum of the light conditions in the area, the size and shape of the target, the density and strength of the target's constituent materials, and the response of the target to a hit (at a given place.) The model is supported by assumptions, conditions, and (presumably,) verification data. (We shall discuss the theory of mathematical duels in this book. The mathematical duel, which we shall simply call a duel in the chapter dealing with mathematical duels, is a special class of war models. Regretfully, there is an ambiguity in the use of the term "duel". Whenever possible, we will use the term formal duel to represent an historical duel; the term duel as a synonym for a mathematical duel.)

We may use this model to build a simulation of a formal duel between two gentlemen. (A formal duel differs from a gunfight in that shots are executed according to a set formula (or model.) Only gentlemen fight duels.) As a simplification, we shall assume that the duel continues until one (or both) of the gunmen is incapacitated or dead. The simulation may be diagrammed as in Figure (I-1).

If we examine this diagram, we notice that random numbers are generated in the simulation to determine the outcome of each exchange of gunfire. Technically, there must be a model in the simulation to generate random numbers of the proper distribution, but such models are not, at this time, germane to this discussion. Further, most computer libraries, and many simulation programs (such as LOTUS 1-2-3, TK! Solver, and MATHCAD,) incorporate one or more random number generators. While random number generators are much used in warfare simulations, their anatomy and physiology are not subjects central to the modeling of warfare. Interested readers should consult a standard text on numerical methods. (e.g. Carnahan, Luther and Wilkes[2]) This simulation can be used to investigate the likely outcome of a formal duel between two gunmen (i.e. the duel may be gamed.) We note again that the simulation incorporates the model(s) in a logical framework of rules that may be used to game a 'real world' event, either past or future, for the purpose of generating understanding.

In passing, we also note that the development of models (modeling) implies the reverse gaming of a simulation.

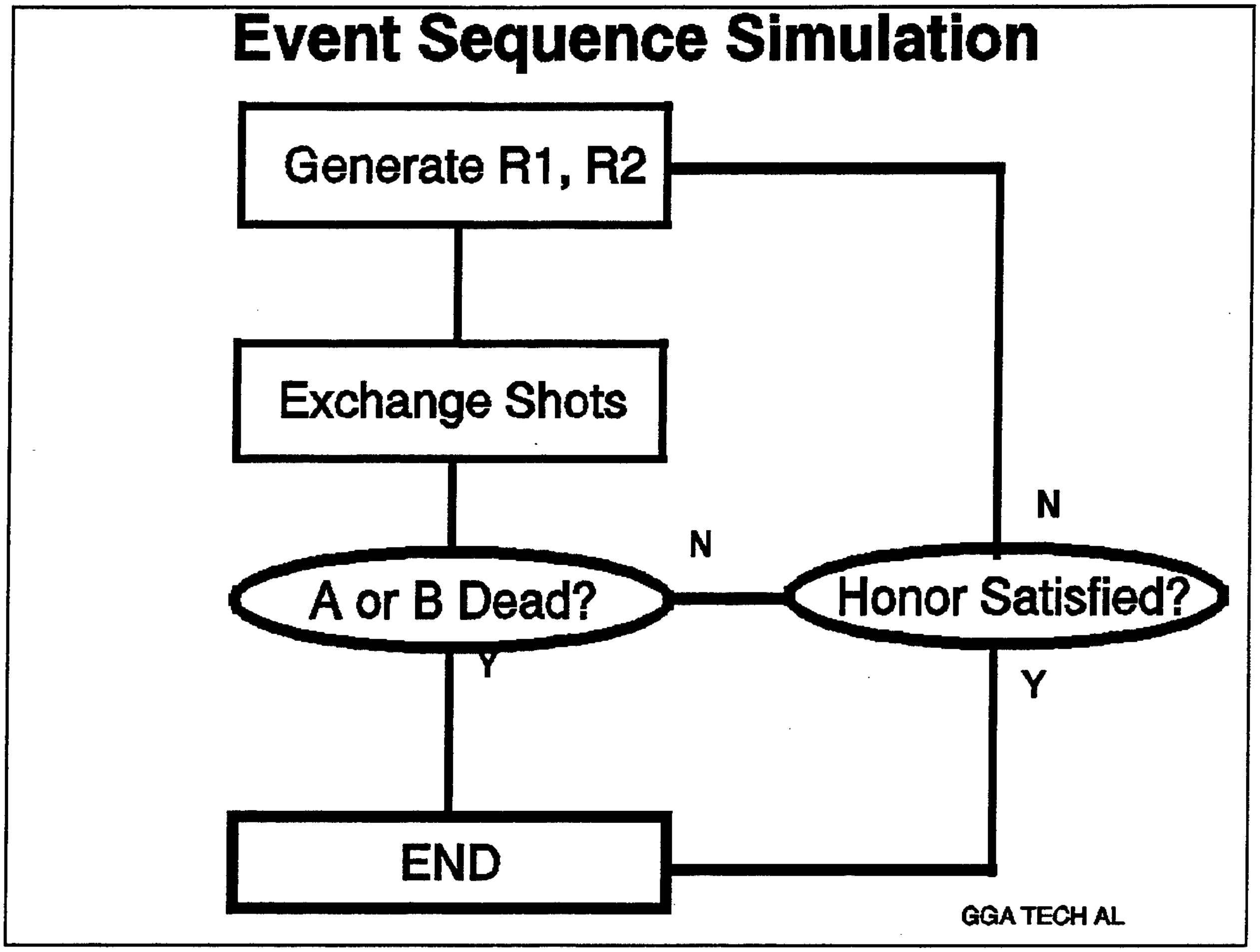


Figure I-1. Duel Simulation

I.D. Types of Simulations

While many authors devote much space to a taxonomy of simulations (or models, depending on their terminology and definitions,[3]) we shall here only briefly describe the different types of warfare simulations.

The most complex of warfare simulations are the iconic, where the model is itself the simulation.

Another type of simulation is the analog simulation; this type of simulation includes parables.

The largest category of types of simulations are the symbolic simulations. These include mechanistic simulations (such as a slide rule,) informational simulations (such as a computer program,) and mixed simulations (such as board games.)

Informational (and mixed) simulations may be either deterministic or stochastic (probabilistic.) Deterministic simulations are those of pure cause and effect; they are not infrequently simulations of the expected values of stochastic models (such as the Lanchester Differential Equations.) Stochastic simulations, also known as Monte Carlo simulations, are usually sequenced by event or time ordering (or, in some cases, both.)

The simulation of the formal duel described in Chapter I.C is an example of an event sequenced simulation. A simulation of a gunfight (where the execution of shots does not occur on a one-to-one, common time start basis,) could be either event or time sequenced (since the shots are not fired simultaneously.) In practice, time and event sequencing are equivalent in philosophy, but care must be taken to ensure that the simulation does not incorrectly favor one side over the other (introduce 'unreal' results) because of the choice of sequencing. This is shown in Figure (I-2). Notice that the sequences of the events are different. This can represent a problem only if the 'real world' is misrepresented. For example, in an event sequenced gunfight

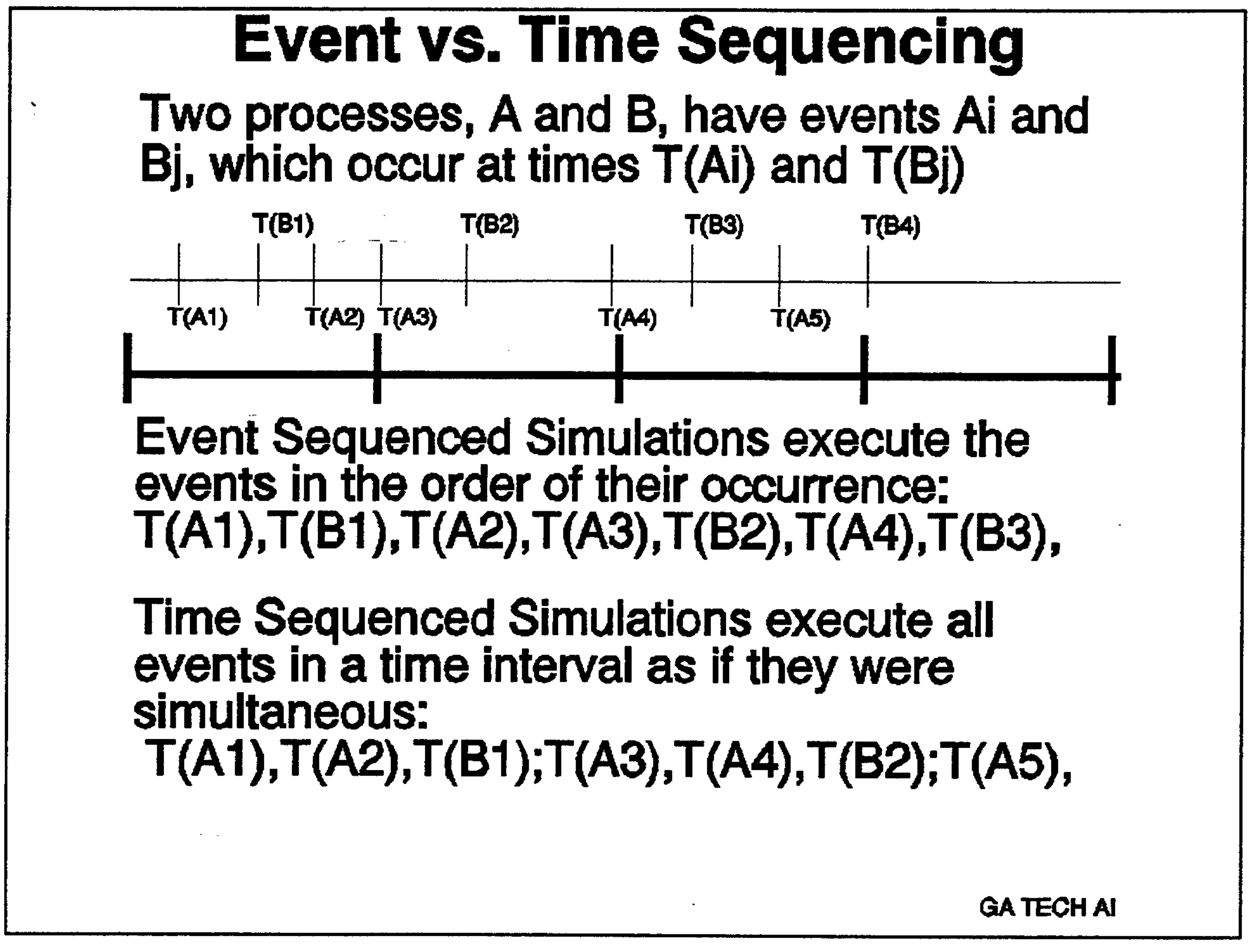


Figure I-2. Sequencing Choices

simulation, a gunman may be allowed to fire after he has been killed. Alternately, in a time sequenced gunfight simulation, the likelihood of both gunfighters dying may be skewed by stopping the simulation too soon (while the killing bullet is in flight. This is an extreme example, but the more complex the simulation, the more likely that such unrealities will creep in.)

I.E. Characteristics of Simulations

Simulations are also often described by characteristics such as the scale of the simulation. Two simulations may emulate combat between two companies of troops, but one simulation may use combat units which are squads while the other simulation may use combat units which are individual troops (or even weapon systems.)

Other characteristics are abstraction versus detail or resolution versus detail. Not all models in the simulation may incorporate the same level of detail in all aspects.

A further consideration is the representation of time and space. Few simulations represent time and/or space continuously. In event sequenced simulations, time dependence may even be hidden or removed. In simulations of only ground troops (at a relatively low level of resolution,) space may be represented only two dimensionally.

Because any simulation must incorporate logical rules to turn itself off, outcome assessment is a characteristic of simulations. As we shall see in the next chapter when we discuss Lanchester's work, a simulation using Lanchester's equations as a model may use a conclusion (total attrition or annihilation of one force) as an outcome assessment.

Finally, simulations may be classified (those used for war gaming at least,) on the basis of how they represent the force. Most conflict simulations are force-on-force; these however may be many-on-many, few-on-few, or one-on-one. Some conflict simulations are one sided (often artillery simulations.) A special class of conflict simulations used in the design of weapons are the engineering simulations.

I.F. War and Simulation

So far, most of our terminology has been fairly general. We now need to get a bit more specific about war. War has many levels, processes and components. Many of these are amenable to modeling and simulation, but not all of them are represented by models in the formal sense. This may seem contradictory, but will become clearer as we continue.

If we take the simple national view of war, then prior to the outbreak of war, there was some opposite non-war state - peace. That state ends in some fashion for

tow (or more) nations and war begins. The common contemporary view, ala Clausewitz, is that this transition results from politics. The causes of war,[4] and even the early stages of war,[5] have been studied and we shall not dwell on them here.[a]

In modern states, there will be standing (i.e., existing) military forces which, depending on the magnitude and duration of the war, may have to be augmented. These forces must be equipped and supplied, trained and moved to the battle, and battles fought. Clausewitz tells us, in a rather dismissive manner, that those actions that are not associated with battles (and their interactions,) are merely preparations for war. War proper is a psychological and physical endeavor for victory.

While there is a great deal of merit to this division, modern experience leads us to believe that war is a national experience that goes far beyond the interaction of military forces on the battlefield. It is useful, however, to drawn an increasing series of distinctions between the processes of "war", and the "preparation for war".

To accomplish this series, it is useful to introduce a hierarchial system of strategic (or war) levels as propounded by Luttwak:[6]

- Grand Strategic,
- Theater,
- Operational,
- Tactical, and
- Technical.

The Grand Strategic level is concerned with the question of war in the large. It is inherently political and economic, as well as military, in nature. The Theater level is concerned with some geographic region where conflict does or may occur. This level serves as a bridge between the highly political Grand Strategy level and the highly military Operational level.

The Operational level is "a middle ground where methods of war contend and battles unfold." The effects at this level are characterized by the contention of armed forces. There are two extremes at this level: attrition and relational maneuver. Neither exists in a pure state. Attritional warfare is the literal grinding away of the enemy's forces, both men and equipment. Relational maneuver warfare (which is basically the same as Liddell Hart's indirect approach,[7]) seeks to incapacitate the enemy by systemic disruption. These two approaches, **attrition** and **disruption**, are fundamental to battle. In contemporary terms, they may be thought of as being "bottoms up" and "top down" approaches to war. Attrition is a bottoms up approach to winning war by wearing down the number of basic military components that make up the military forces. Relational maneuver is a top down approach to prevent the enemy from using his force effectively and decisively.

---

[a] We shall consider several models associated with the non-combat aspects of war in later chapters.

Both of these approaches are organizationally oriented. Military forces are not simple collections of men and equipment. These forces have an organizational structure to make them effective. Denial of that effectiveness has two fundamental forms - attrition and disruption. Attrition reduces the effectiveness of a military organization through its component parts. Disruption reduces the effectiveness at the organizational level itself.

The tactical level of war (in Luttwak's hierarchy,) is primarily concerned with the battlefield interactions of these organizations (admittedly at a low level,) and their components. The technical level is primarily concerned with the interactions of the components, usually on a one-on-one basis, and is largely dominated by matters of physics, engineering, and doctrine. The majority of this book is concerned with discussing some analytical tools and techniques for describing these two levels of war.

At the Grand Strategy level of war, the primary concern is political and economic. Questions concerning the production (and development) of war materials, their transport, and the recruitment, training, and transport of troops are amenable to modeling and simulation. Indeed, there are extensive simulations of these process in place. Additionally, there are political and national will/morale processes, which, being human dominated, are less amenable to modeling and simulation. At this level, the interactions of the military forces have importance primarily as they effect these latter processes although there are exceptions such as the tactical questions of convoy attrition on supply.[b]

As we proceed down the hierarchy, the logistical questions of supply, transport and training become less important and the interactions of the military forces become more important. This trend culminates at the tactical and technical levels where the availability of men and materials, their training state and the nature of tactical doctrine become essentially parametric.

Except at the technical level, which is dominated by physical processes, all of these levels have considerable psychological or human components. At the tactical level, morale and willingness to fight (or surrender,) are potentially important factors, and the reader must be warned that it is in this psychological area that models and simulations of war, at whatever level, are most primitive and *ad hoc*. As we shall see, the tactical models that we discuss do not, in and of themselves, consider termination of combat except in limiting mathematical form.

---

[b] The convoy question is a classic of military operations research. Although we are concerned here with the Grand Strategy level of war, the effective transportation of men and materiel across oceans has an operational/tactical component concerned with protecting the transporting vessels from attrition.

Admittedly, some of the technical models that we will discuss do have human components, but these components are essentially physiological rather than psychological in nature. In general, the psychological aspects and processes of war are those which are the least represented in combat simulations.[c]

I.G. Combat and Simulations

As we have previously stated, much of this book is concerned with the tactical and technical levels of war. Our central focus will be on models of combat processes. Although we shall speak in general of combat, most of our discussion will be concerned with duels[d] or engagements, and battles.[e] To a slight extent, we shall be concerned with the operational level of war which is primarily concerned with campaigns. Campaigns may be thought of as an orchestrated (hopefully) series of battles. What raises this series above the tactical level is the non-combat processes which occur (e.g., relational maneuver.)

At the technical and tactical levels of war, the primary modeling interest is the interactions between the individual weapon systems. One-on-one engagements are usually considered to be technical for Army weapons, but tactical for Air Force and Navy weapons. This can be seen easily by considering that most warships and warplanes carry more than one weapon system.

The interactions between two weapon systems (with crew,) or between a weapon system with crew and a man (men) or vehicle are probabilistic in nature, That is, if an infantryman fires his rifle at an enemy rifleman, there is some probability that a hit will occur. This probabilistic nature is fundamental to our approach to combat processes, and largely determines the two approaches to the modeling and simulation of combat.[f]

---

[c] An actual argument may be made that there are social processes in the simulation community that act to prevent development and inclusion of psychological models.

[d] Duels, as used hereafter, have a fairly rigorous mathematical formulation. They are considerably more general than our picture here of two men shooting at each other. We shall briefly review duel theory in a later chapter.

[e] We will make a somewhat confusing use of the terms engagement and battle. We shall use the term engagement to mean both the firing of a weapon at a target, and the overall combat processes in combat between two or more forces. The term battle will tend to be reserved for one or more engagements (second meaning) and possibly maneuver, reinforcement, resupply, etc.

[f] This does not mean that there are combat processes or subprocesses that are actually or practically deterministic. They do exist, as we shall see. In a general sense, however, they are the exception rather than the rule.

These two basic approaches are **specification** and **aggregation** of the forces. The specification approach is more obviously probabilistic in nature. For this reason, simulations built using this approach are commonly referred to as Monte Carlo or Stochastic simulations. Under this approach, each individual weapon system/platform is represented explicitly. Combat processes are represented individually and probabilistically. That is, each time an infantryman fires at an enemy, the simulation generates a random number to determine whether a hit occurs or not. A key technical concern of these simulations is thus the generation of these random numbers. Because of the individual representation of each weapon system, these simulations tend to be quite large in size and may require multiple executions to arrive at statistically significant results.

The aggregation approach lumps together weapon systems. Forces are usually represented by their strengths (numbers.) Different types of aggregation may be used on the basis of weapon system and/or organization type. Combat processes are represented by rate of change of force strength. These rates are usually the expected values (and occasionally the standard deviations,) of the relevant combat processes. For this reason, simulations based on aggregation are often referred to as deterministic or expected value. These simulations may be smaller in size than specification simulations, and generally do not require multiple executions. Key technical concerns with aggregation simulations are how the rates are calculated from the combat process models, and the form and technique of solution of the force strength relations.

In general, the same probabilistic combat process models go into both types of simulations. The two types of simulations differ in how the expected values (and standard deviations,) of the processes which make up the entire battle are calculated. In Monte Carlo simulations, the expected values for the battle are calculated by the simulation. In aggregated simulations, the expected values of individual engagements are calculated before the simulation is executed. There are several significant differences between these two types of simulations, but the most important commonality is that they share combat process models. Thus, if we know these models, we know what goes into each type of simulation.

The aggregation models of the changes in force strength are the philosophical basis for the aggregated simulations, and arguably, for the Monte Carlo simulations as well. From a theoretical standpoint, these models are essentially all attrition models - the other combat processes do not have the theoretical framework that attrition does!

The theory for these models was initiated at the start of the Twentieth Century by four men. One of them, Lanchester, is generally credited with the basic work although, as we shall see, the question of who was the founder of attrition theory is largely moot. The next chapter reviews the lives and contributions of these four

pioneers.

References

1. Callahan, Leslie G., "Modeling, Simulation, and Gaming of Warfare", Twelfth Annual Course on Modeling, Simulation and Gaming of Warfare, School of Industrial and Systems Engineering, Georgia Tech Continuing Education, Atlanta, GA, 13-16 August 1991.

2. Carnahan, Bruce, H. A. Luther, and James O. Wilkes, **Applied Numerical Methods**, John Wiley and Sons, New York, 1969.

3. Hughes, Wayne P., Jr., ed., **Military Modeling**, Military Operations Research Society, Alexandria, VA, 1984.

4. Wright, Quincy, **A Study of War**, University of Chicago Press, Chicago, IL, 1942.

5. Ivanov, S., **The Initial Period of War**, Moscow, 1974, Soviet Military Thought Series No. 20.

6. Luttwak, Edward N., **Strategy: The Logic of War and Peace**, Harvard University Press, Cambridge, 1987.

7. Liddell Hart, B. H., **Strategy**, Frederick A. Praeger, New York, 1967.

II. LITERATURE REVIEW I

II.A. Introduction

This chapter is the first of several which review the literature of the dynamics of warfare. This chapter is devoted to the origins of that literature. As such, it is devoted to the work of the four founders of the discipline: Lanchester, Osipov, Fiske, and Chase; an Englishman, a Russian, and two Americans.

II.B. Frederick William Lanchester

Who is this man whose name is uniquely associated with the dynamics of warfare? What little[1] we know indicates that Lanchester was a Research and Development engineer of great accomplishment, a pioneer in the development of automobiles, aircraft, and operations research. The latter is of primary interest here.

In retrospect and in the context of our own day, it seems obvious now that Lanchester continuously sought out problems and solved them, but was not greatly concerned with turning his solutions to practical applications. In this he typifies the developing discipline of bellum dynamics; he pursued (what are now) academic problems without the benefit of an academic environment. As a result, many of his accomplishments found no recognition until years later; indeed, his efforts consistently verged on the edge of failure because of their non-application to Civilization's affairs. His work on warfare dynamics typifies this; performed during World War I, it found little or no application during World War II and recognition only in the years following that war.[a]

This is not to portray that the man was a failure. Rather, he was draped in most of the scientific honors that Imperial England could bestow save only knighthood. Only in the area of economic success could Lanchester be reckoned a failure, especially in the automobile industry.

It is, however, in the field of operations research that he has become a demigod, immortalized in the ultimate award of that field.

The seminal contribution of Lanchester to operations research is contained in

---

[a] Until recently, it was commonly believed that Lanchester's theories were not widely known. However, the recently (1988?) discovered correspondence of (then) Captain J. V. Chase, USN, of 1921 indicates, at least, his, and presumably, his correspondent's familiarity with the concepts. It seems reasonable that the work of Lanchester enjoyed some attention within the officer corps of the U.S. Navy prior to World War II. It was only after that war, possibly due to the development of the digital computer, that an industry based on Lanchester's work came into being.

his book **Aircraft in Warfare: The Dawn of the Fourth Arm** published in 1916,[2] now sadly out of print and difficult to find in any but the most venerable of libraries. (Indeed, the copy that I was able to study was provided under interlibrary loan from the United States Military Academy library.)

This work runs to nineteen chapters, of which V, VI, and XVIII are most relevant to the subject at hand. The following text presents a brief outline of the material contained in those chapters, as it applies to Bellum Mechanics.

The Principle of Concentration which underlies what would become known as the Lanchester Equations begins Chapter V, dated October 21, 1914.

> **The Principle of Concentration: the force with the greater strength, all other factors being equal, will inflict the greater damage.**

This principle is illustrated by the difference between ancient and modern warfare. In ancient warfare, combat is typified by an essentially linear interaction of troops at a combat interface engaged in a one-on-one (short range) manner. (The use of long range weapons such as the crossbow or the longbow is conveniently ignored.) In modern warfare, combat is typified by a more areal interaction of troops in a many-on-many (long range) manner.

The Principle of Concentration leads directly to the definition of the Quadratic Lanchester Differential Equation:

> **"If, we assume equal individual fighting value, and the combatants otherwise (as to 'cover,' etc.) on terms of equality, each man will in a given time score, on the average, a certain number of hits that are effective; consequently, the number of men knocked out per unit time will be directly proportional to the numeric strength of the opposing force. Putting this in mathematical language, and employing symbol b to represent the numerical strength of the 'Blue' force and r for the 'Red', we have:**

$$\frac{db}{dt} = -r\,c \;.....(1) \qquad \text{(II.B-1)}$$

> **and**

$$\frac{dr}{dt} = -b\,k \;.....(2) \qquad \text{(II.B-2)}$$

> **in which t is time and c and k are constants (c = k if the fighting values of the individual forces are equal)."**

{In contemporary terms, the constants c and k are called attrition or kill rates. They will be designated throughout this book by Greek letters, usually α and ß.}

The Principle of Concentration is illustrated by an example: Consider two forces of 1000 men each. The red force is divided into two units of 500 men each which serially engage the single (1000 man) unit of the blue force. {Lanchester introduces his differential equations and the state solutions to them, but not the explicit time solutions. We shall develop these in Chapter III, but we introduce here the state solution of the Quadratic Lanchester Differential Equation:

$$B^2 - B_0^2 = \frac{\alpha}{\beta} (A^2 - A_0^2) \qquad \text{(II.B-3)}$$

where: B, A are the force strengths of the 'Blue' and 'Red' (Amber) forces, respectively,

B_0, A_0 are the initial force strengths of the 'Blue' and 'Red' forces, respectively, and

α, ß are the attrition rates (kills per unit time per man) for the 'Blue' force against the 'Red' force, and the 'Red' force against the 'Blue' force, respectively.

If we take the attrition rates to be equal, then the two serial combats may be modelled:

First Engagement	
$B_0 = 1000$	$A_0 = 500$
$B = 866$	$A = 0$
Second Engagement	
$B_0 = 866$	$A_0 = 500$
$B = 707$	$A = 0$

This example, which depicts the Blue force totally destroying the Red force (100% loss) with only moderate loss (30%) to itself by being able to concentrate, illustrates the Principle of Concentration and supports the axiom of war that forces are not to be divided.

It seems equally obvious that this example is for illustrative purposes only. Battles do not proceed (usually - we discuss this in a later chapter) to the complete destruction of one force (which Lanchester calls a conclusion.) Unfortunately, Lanchester introduces, almost immediately, this mathematical concept of victory prediction as complete attrition of one force - the concept remains with us to this day.

In his book, Lanchester presents a graph to depict the general weakness of a

divided force. Rather than reproduce the graph here, we instead take an algebraic approach. Again let the initial force strengths of B and A be the same, and let the attrition rates for the two forces be equal. We rewrite the state solution, Equation (II.B-3), in the form

$$B^2 - B_0^2 = A^2 - A_0^2. \tag{II.B-4}$$

(We have dropped the attrition rates α and ß since they are equal and cancel - the ratio α/ß has a value of one.)

If the battle is fought through to a conclusion, then one of the two forces is completely attrited. Since the initial force strengths are givens, the 'victor' is entirely decided by the sign of the right hand side of this equation. If the right hand side is positive, the Blue force is the victor (or survivor); if negative, the Red force is the victor; if zero, a draw occurs (which presumably ends in mutual destruction!)

Since the combat occurs serially, we may write the initial Red force strength squared as

$$A_0^2 = N^2 X^2 + N^2 (1 - X)^2, \tag{II.B-5}$$

where N is the initial total strength of the Red force, and X is the fraction of the Red force in the first unit.

Since the initial strength of the Blue force is also N, the state solution, Equation (II.B.-4) may be rewritten, using Equation (II.B-5), as

$$B^2 - A^2 = N^2 (2X - 2X^2). \tag{II.B-6}$$

We may immediately see from Equation (II.B-6) that any division of the Red force results in a Blue force victory (assuming the combat is carried to a conclusion) since

$$2X - 2X^2 > 0, X < 1, \tag{II.B-7}$$

Only if X = 1 (an undivided Red force) does the combat become a draw.

We may further see that Blue force losses are minimized when X = 0.5 (an even division.) This example addresses only the case where the two forces and their attrition rates are equal. The Quadratic Lanchester Law - Principle of Concentration can be used to develop cases which predict an advantage for the division of forces. An example of this would be a division of the Red force where part of the force is used to execute a flank attack on the Blue force. (The combat is now not serial, but staggered.) In this case, it is convenient to write the state solution in the form

$$B^2 - \frac{\beta}{\alpha} A^2 = B_0^2 - \frac{\beta}{\alpha} A_0^2. \qquad \text{(II.B-8)}$$

The initial Red force strength squared, times the ratio of attrition rates, is

$$\frac{\beta}{\alpha} A_0^2 = \frac{\beta}{\alpha_1} N^2 X^2 + \frac{\beta}{\alpha_2} N^2 (1 - X)^2, \qquad \text{(II.B-9)}$$

where α_1, α_2 are the attrition rates against the frontal attack and flank attack units, respectively. (Note that the attrition rate of the Red force against the Blue force is the same for both the frontal and flank attack units (i.e. ß); it is the attrition rate(s) of the Blue force against the Red force that changes with force engaged.)

If, for convenience, we assume that the attrition rates for frontal attack are the same for both forces, and that the forces again have the same initial force strength, then we may write

$$B^2 - A^2 = N^2 (1 - X^2) - \frac{\beta}{\alpha_2} N^2 (1 - X)^2. \qquad \text{(II.B-10)}$$

This equation shows that (for example,) if 10% of the Red force is put into the flank attack, and if the vulnerability of that force to attrition by the Blue force is reduced, through surprise or whatever, by a factor of at least 20, then the victory will be Red's rather than Blue's.

Lanchester is apologist in defending the validity of counting the numbers which comprise the forces on the grounds that the counting will be done anyway. He further asserts that training and morale are not suited to theoretical discussion, the performance of weapons is. The use of weapons in combat is dependent on the morale and training of the troops. If the troops are not trained, they cannot use their weapons. Nor, if their morale suffers, are they likely to use their weapons. It is not that these have no effect; rather, Lanchester asserts, they are not amenable to analysis. The question of what constitutes the strength of a unit is best expressed by two quotes:

> **"The fighting strengths of the two forces are equal when the square of the numeric strength, multiplied by the fighting value of the individual units, are equal."**

and

> **"The fighting strength of a force may be broadly defined as proportional to the square of its numerical strength multiplied by the fighting value of its individual units."**

This is basically the same as we demonstrated in the previous example for adding forces.

If the attrition rate of a machine gun is 16 times that of a rifleman, then 250 machine guns (with crews) have the force strength of 1000 riflemen. In an engagement between a force of riflemen and a force of machine guns, the individual machine gun will (on the average) receive four times as much fire as an individual rifleman would under the circumstances. This is true if the fire of the riflemen is aimed (as in the Boer War - Lanchester's example.)

If on the other hand, fire is distributed without such pinpoint aiming over the area covered by the force, then the machine gun will receive only slightly greater fire than a single rifleman would (given a slightly larger area for the machine gun,) and may actually be less. For example, say that both forces hold an area of 10 square kilometers. This equates to 1,000 m^2 per rifleman or 4,000 m^2 per machine gun. Both are subjected to fire from an area weapon with an area of effectiveness of 100 m^2. This translates into an attrition rate of 0.1 rifleman per fire, but only 0.025 machine guns per fire.

This line of reasoning leads to the Linear Lanchester Differential Equations:

$$\frac{dA}{dt} = -\alpha\ A\ B, \qquad \frac{dB}{dt} = -\beta\ B\ A, \tag{II.B-11}$$

where A and B are the force strengths of the Red and Blue forces respectively, and α and ß are the attrition rates. {Note that these α and ß are different from the previous ones for the Quadratic Lanchester Differential Equation.} Lanchester notes that in this case where fire is directed against an area and not against an individual, the rate of loss is independent of numbers and dependent only on the efficiency of the weapons. In this case, there is no value in concentration. This case is cited as being more appropriate for describing ancient combat, not because the weapons are long range, but because the units were only engaged along a linear interface and thus the numbers engaged at any moment, on either side, were approximately equal. We may note however, that in modern terms, the Linear Lanchester Differential Equations are normally appropriate for describing the use of what the Russians call Weapons of Mass Destruction, in particular nuclear and chemical weapons. (Whether they are appropriate for biological weapons depends on the exact vector(s). We shall comment further on this in a later chapter.)

It is interesting to note that even in ancient combat, there appears to be advantage in concentration in the line. Notable ancient success stories such as the Greek phalanx and the Roman legion enjoyed considerable increase in force strength by, in effect, concentrating more men into the linear interface. Notably, this was the result of better tactics, training, doctrine, and/or morale, which Lanchester states are not amenable to analysis.

The principle historical analysis presented by Lanchester to illustrate the Quadratic Law and the Principle of Concentration is the Battle of Trafalgar (1805). Here the British fleet (40 ships) under Admiral Lord Nelson divided the French Fleet (46 ships) and engaged the rear half at a force ratio of 32:23. This gave a total force strength of

British	French
$(32)^2 + (8)^2$	$(23)^2 + (23)^2$
1088	1058

which should, if carried to a conclusion, have resulted in a draw.

The dynamics of ship motion were such that a significant period of time would be required for the front half of the French fleet to sail back to the aid of its rear half. Further, Nelson used 8 of his ships to slow this process. Thus, in the main battle area, the force strength ratio was

British	French
$(32)^2$	$(23)^2$
1024	529

Which gives a force strength ratio, British to French of approximately 2:1. This analysis is based on an operational memorandum prepared by Nelson before the battle and the actual forces are somewhat different. The outcome was not.

1. McCloskey, Joseph F., "Of Horseless Carriages, Flying Machines, and Operations Research", Operations Research, 4 141-147, 1956.

2. Lanchester, Frederick W., **Aircraft in Warfare: The Dawn of the Fourth Arm**, Constable and Company, LTD., London, 1916.

II.C. Osipov

In their book Forecasting in Military Affairs, Chuyev and Mikhaylov[1] devote approximately two pages (out of 230) to the area of Lanchester's equation as partial contents of a section on Differential Equation Models in a chapter on Military Formalism models. Most of that two pages of text consists of two example: one is concerned with the Quadratic Lanchester Differential Equations and their state solution, and the other is concerned with a transport theory outgrowth (which we shall treat in a later chapter.) No mention is made of the Linear Lanchester Differential Equations, nor does the example present any numeric data (a relative rarity among Russian authors who seem enamored with including numerous tables of data in their works -perhaps an indication of the lack of computational capability avail-able to the student? Or a potential embarrassment to the state since historic casualty data would be needed?)

What is most startling in these two pages is the claim that the "Lanchester equations" had been put forward by Osipov earlier. No reference to this work by Osipov is given.

Searches for the work proved fruitless, given the paucity of real information. For a while, I ascribed Osipov to be another piece of Russian hype, claiming the development of Lanchester's equations just as they had similarly claimed to have invented everything from the Franklin stove to the fundamental theorem of the calculus. Subsequently this ascription proved to be false. The Library of Congress yielded up to Dr. Allan Rehm five articles by one M. Osipov[2] all published during 1915. Subsequently, Dr. Rehm advised me that two separate translations had been made, one by Dr. Helmbold and another by Deborah Coulter-Harris[3] of the Soviet Army Studies Office at Ft. Leavenworth. He was kind enough to provide me with a copy of the latter. The remainder of this section is based on that translation. (Where there are "direct" quotes, they should be taken in the context of the translation. Subsequently, Drs. Helmbold and Rehm have made their own translation.[4])

It is clear from Osipov's articles that he developed his theory of combat independently of Lanchester and Fiske. Not only are the tone and texture of the material different, but there is significant new material and philosophy. Further, despite his protestations that he is (was) neither a specialist in military history nor skilled in the practice of military matters, he is, manifestly by his knowledge and arguments, a student of both. He also has a knowledge of mathematics and statistics, although it also seems unlikely that he is either a professional mathematician or statistician. His ability to communicate in writing is evident, even in English translation, yet his antipathy to the press would seem to indicate that he is not a journalist. If we proceed with this fanciful analysis, we would be led to speculate that M. Osipov is a teacher (this would explain his communication skills and his well rounded, but apparently introductory knowledge of history and mathematics,) who

served in the army during the Great War.

Osipov begins his article by considering history. He almost immediately provides a list of 38 battles spanning the century from 1805 to 1905. This list excludes battles between "regular troops and disorganized elements of uncivilized countries" (colonial battles,) and "battles where one side has a fortress or strong temporary fortifications." While he does not consider the duration of the battles, he does list initial force strengths and losses (equivalent to final force strengths) for both sides. These are organized by stronger force versus weaker force (initially) rather than by victor-loser or attacker-defender. However, of the 38 battles, 28 were won by the stronger side. These are shown in Table (II.1).

Just as Lanchester introduced the Law of Concentration, Osipov introduces the Law of Distribution of Losses (or just Law of Losses):

> **"Law of Distribution of Loses: The strongest side has less losses than the weaker side."**

If we take (in our preceding notation) Red to be the stronger side and Blue to be the weaker, then we may write this mathematically as

$$A_0 - A < B_0 - B. \tag{II.C-1}$$

We immediately see an apparent conflict between Lanchester and Osipov since the Law of Distribution of Losses states that Lanchester's Quadratic Law does not hold. This however, is not the case if we consider the role of the attrition rate constant/-function. If we compare this equation to Lanchester's linear law state solution, we find that the Law of Distribution of Losses gives us the requirement that

$$\frac{\alpha}{\beta} < 1, \tag{II.C-2}$$

for the Linear law. For the Square law, a somewhat more complicated situation exists. To investigate this, it is convenient to write the square law state solution in the form,

$$\beta (A_0 + A)(A_0 - A) = \alpha (B_0 + B)(B_0 - B), \tag{II.C-3}$$

which we may rewrite as

$$(A_0 - A) = \frac{\alpha (B_0 + B)}{\beta (A_0 + A)} (B_0 - B), \tag{II.C-4}$$

since the right hand side is, by the Law of Distribution of Losses, less than blue's

losses (B_0 - B), this equation reduces to

$$\alpha (B_0 + B) < \beta (A_0 + A), \tag{II.C-5}$$

or

$$\frac{\alpha}{\beta} < \frac{(A_0 + A)}{(B_0 + B)}. \tag{II.C-6}$$

We may define the losses as a $\equiv$ (A_0 - A) and b $\equiv$ (B_0 -B), which allows us to rewrite Equation (II.C-6) as

$$\frac{\alpha}{\beta} < \frac{(2A_0 - a)}{(2B_0 - b)}. \tag{II.C-7}$$

Since $A_0 > B_0$ by postulate (and convention), and a < b by the Law of Distribution of Losses, it immediately follows that

$$(2A_0 - a) > (2B_0 - b), \tag{II.C-8}$$

so that the ratio α/β is less than some number greater than one for Lanchester's square law. Thus, there is no conflict between Lanchester and Osipov on the basis of the mathematical formulation of the Law of Concentration and the Law of Distribution of Losses. It remains to be seen if this is supported by historical evidence.

This historical evidence is one of the primary contributions of Osipov in his articles. As we have stated, Osipov presents a table of 38 battles. The dates of these battles span the century 1805-1905. Most are drawn from the Napoleonic era (1805-1815) or during the thirty year period 1850-1870. (Crimean War, Second War of Italian Independence, Austro-Prussian War, and Franco-Prussian War). (We shall examine these data in greater detail in a later chapter devoted to historical insights. Our comments in this chapter will be limited to a review of Osipov's five articles).

Of these 38 battles, Osipov notes that 28 were victorious for the force with the greater numbers. (We note that Osipov rounds all of his numbers, usually to thousands. This gives rise to some calculations which appear more definite in their significance than if rounding had not been performed. This is especially true in the statistical inferences that Osipov draws). From a companion of losses in these battles, Osipov concludes that, in general, the stronger side takes fewer losses than the weaker side in a battle. He quickly notes, however, that there are many other factors which influence the outcome of the battles. What is significant in that in the

consideration of the battles as an aggregate data set, the effects of these factors are decreased and the effects of pure numbers may be seen. (This is one of Osipov's significant new contributions. By aggregating these battle data, he in essence takes a scientific approach to the problem, asserting that factors other than pure numeric strengths may be treated as random error sources (relative to the calculation at hand), which cancel out in the mean).

Osipov next describes a "simplest method of calculating losses" which is to all intents and purposes is Lanchester's Quadratic Attrition differential equation. He writes the state solution as

$$A'^2 - A_0^2 = B'^2 - B_0 2, \tag{II.C-9}$$

where: $A' \equiv A_0 - a$
$B' \equiv B_0 - b.$

He further advances the approximation

$$A_0 \, a = B_0 \, b, \tag{II.C-10}$$

based on examination of his table of historical data. (In Lanchestrian terms, Osipov is stating that history indicates that most battles stop far from a conclusion. We will examine this in more detail in a later chapter).

Next, using a calculus argument, and introducing an attrition rate (identical for both sides), Osipov derives the analytic Quadratic Law solutions as a function of time. Tables of the cosh and sinh functions are presented since they would not normally (?) be available to the reader, and example calculations are presented. Osipov then notes that the time solutions are "not appropriate for application to military history, because α (attrition rate) and t (time) are unknown". This statement recognizes two fundamental problems in the analysis of historical data:

* how to get battle duration data, (a very difficult undertaking), and
* how to use it why you have it, since combat is not continuous.

(Again, we shall treat this in more detail in the chapter on historical insight). In the process of developing further examples, Osipov presents the Quadratic Law solutions for distinct (i.e. different) attrition rates, but claims that the derivation is so similar to the previous one that he will not take the space to belabor it. He does, however, present the "modified state solution, Equation (II.B-3).

Osipov next introduces consideration of a force comprised of two different weapon systems (rifles and some other weapon such as machine guns or direct fire artillery). With the assumption that the second type of weapon takes no casualties (is not atritted) he writes the Quadratic solutions as

$$\left(A + \frac{\beta}{\alpha}M_0\right) = \left(A_0 + \frac{\beta}{\alpha}M_0\right)\cosh(\alpha t) - \left(B_0 + \frac{\beta}{\alpha}N_0\right)\sinh(\alpha t) \tag{II.C-11}$$
$$\left(B + \frac{\beta}{\alpha}N_0\right) = \left(B_0 + \frac{\beta}{\alpha}N_0\right)\cosh(\alpha t) - \left(A_0 + \frac{\beta}{\alpha}M_0\right)\sinh(\alpha t),$$

where: β is the attrition rate for the second type of weapon (e.g. machine guns)
α is the attrition rate for the first type of weapon (rifles)
A_0, B_0 are the initial number of rifle bearing troops (assumed one to one) for each side respectively, and,
M_0, N_0 are the initial (constant) number of weapons of the second type.

The state solution is

$$\left(A + \frac{\beta}{\alpha}M_0\right)^2 - \left(A_0 + \frac{\beta}{\alpha}M_0\right)^2 = \left(B + \frac{\beta}{\alpha}N_0\right)^2 - \left(B_0 + \frac{\beta}{\alpha}N_0\right)^2. \tag{II.C-12}$$

Note that Osipov states that this formalism is valid only if the number of second type of weapons is not atritted.

Osipov goes on to state that this technique of normalizing the attrition of additional weapons may be extended to third, fourth, etc. types of weapon systems so long as they are not atritted. He also expands the state solution for small losses as

$$(A_0^2 - A^2) - (B_0^2 - B^2) + 2\frac{\beta}{\alpha}(aM_0 - bN_0) = 0. \tag{II.C-13}$$

He then proceeds to calculate the ratio β/α, the relative attrition of artillery (in this case) to rifles and finds it is a number ~ 123 -143, for these Napoleonic battles.

One of the concepts Osipov introduces is the "correlation of losses". He compares the actual losses to a calculation based on the other strength numbers. While there is no basis for the association, it is still interesting to postulate that this type of calculation maybe the genesis of the "Correlation of Forces" practiced in the Russian armed forces today. Certainly, a logical connection can be made between the types and forms of the calculations.

Osipov introduces the differential equation

$$\sqrt{A}dA = \sqrt{B}dB, \tag{II.C-14}$$

based on his analysis of the historical data. This gives rise to the state solution

$$A^{3/2} - A_0^{3/2} = B^{3/2} - B_0^{3/2}. \tag{II.C-15}$$

In the correlation of losses calculation, Osipov calculates the losses for the stronger sides using the state solutions for the Quadratic Law (from Equations (II.B-3) and (II.C-10), respectfully) and the 3/2 law (Equation (II.C-15) above). The difference between this calculated loss and the actual loss is treated as an error term and the aggregate for the 38 battles is treated to an error analysis. (Osipov rounds to thousands here, resulting in a tidier result than would be found otherwise. The numbers in the Table II-2 are not rounded due to the way that the table was formed). He finds average errors of 22% for the exact quadratic law state solution, 15% for the approximate quadratic law state solution and 0.7% for the approximate 3/2 law state solution. Further, the mean is essentially the median for the 3/2 law calculation. He concludes that the 3/2 law most clearly describes this type of battle.

He further concludes that for small battles (<75,000), the quadratic law may be more relevant than the 3/2 law. For force strengths >75,000, the 3/2 law appears to be more relevant. Osipov does note, however, that the rationale for the 3/2 law is purely empirical while the quadratic law is better founded theoretically. (We shall examine the 3/2 law in greater detail in the chapter on Osipovian combat. Since Osipov and Lanchester appear to have independently developed mathematic attrition theories with many points in common, we shall adopt the following nomenclature: The quadratic and linear attrition processes will continue to be referred to as Lanchestrian rather than as the more cumbersome Lanchestrian - Osipovian, attrition process other than quadratic and linear, will be termed Osipovian in recognition of the greater generality of Osipov's empirical consideration of attrition).

Osipov next proceeds to consider further the statistical aspects of his theory. He examines error sources such as leadership, morale, reserves, artillery, weapons quality on terrain and improvements, large number of fighting units, density of fighting units and (considered to be systematic errors). He examines the concept that battles terminate when one side has taken 20% losses.

Osipov concludes by stating that the dependance of losses on the numerical strength of the forces exists but cannot be verified except on a statistical basis. However, the stronger side has less losses than the weaker side. (Law of distribution of losses). He does not present his theory as other than an example of the application of existing military principles.

It seems likely that Osipov's papers were not all that well received by the Russian media when they were published. Certainly we do not know what happened to Osipov following their publication. We do know that they have been used, in some form, in the Military Operations Research community of the USSR.

1. Chuyev, Yu. V., and Yu. B. Mikhaylov, **Forecasting in Military Affairs**, Moscow, 1975, Volume 16 in Soviet Military Thought, U.S. Government Printing Office, Washington.

2. Osipov, M.,"The Effect of the Quantitative Strength of Fighting-Sides on the Losses", Voennie Shornik, 3-7, 1915.

3. Coulter-Harris, Doborah, "Translation of The Effect of the Quantitative Strength of Fighting-Sides on the Losses", Soviet Army Studies Office, Ft. Leavenworth, KS, 1987.

4. Helmbold, Robert L., and Allan S. Rehm, trans., M. Osipov, "The Influence of the Numerical Strength of ENgaged Forces on their Casualties", U.S. Army Concepts Analysis Agancy, Bethesda, MD, Research Paper CAA-RP-91-2, September 1991.

Battle	Stronger Force	Start	Losses	Weaker Force	Start	Losses	Date
Austerlitz	Allies	83	27	French	75	12	1805
Jena	French	74	4	Prussians	43	12	1806
Auershtedt	Prussians	48	8	French	30	7	1806
Preisish	French	80	25	Russians	64	26	1807
Freiland	French	85	12	Russians	60	15	1807
Aspern	Austrians	75	25	French	70	35	1809
Wagram	French	160	25	Austrians	124	25	1809
Borodino	French	130	35	Russians	103	40	1812
Berezina	Russians	75	6	French	45	15	1812
Bautsen	French	163	18	Allies	96	12	1813
Ganau	French	75	15	Allies	50	9	1813
Drezden	Allies	160	20	French	125	15	1813
Keiptsig	Allies	300	50	French	200	60	1813
Katsbach	Allies	75	3	French	65	12	1813
Liutsen	French	157	15	Allies	92	12	1813
Dennevits	French	70	9	Allies	57	9	1813
Kul'm	Allies	46	9	French	35	10	1813
Laon	Allies	100	2	French	45	6	1814
Kpaon	French	30	18	Russians	18	5	1814
Waterloo	Allies	100	22	French	72	32	1815
Lun'i	French	120	11	Prussians	85	11	1815
Grokhoro	Russians	72	9	Poles	56	12	1831
Al'ma	Allies	62	3	Russians	34	6	1854
Chernaia	Allies	62	2	Russians	56	8	1854
Inkerman	Russian	90	12	Allies	63	3	1854
Col'ferino	Austrian	170	20	French	150	18	1859
Madzhenta	Austrians	58	10	French	54	5	1859
Kustotsa	Austrians	70	8	Italians	51	8	1866

Battle	Stronger Force	Start	Losses	Weaker Force	Start	Losses	Date
Kenigrets	Prussians	222	10	Austrians	215	43	1866
Mets	Germans	200	6	French	173	20	1870
Gravelot	Germans	220	20	French	130	12	1870
Mars LaTour	French	125	16	Germans	65	16	1870
Vert	German	100	10	French	45	5	1870
Sedan	Germans	245	9	French	124	17	1870
Aladzha	Russians	60	2	Turks	36	15	1877
Shabh	Russians	212	40	Japanese	157	20	1904
Liaoian	Russians	150	18	Japanese	120	24	1904
Mukden	Russians	300	59	Japanese	280	70	1905

Battle	Stronger Force Start	Stronger Force Losses	Weaker Force Start	Weaker Force Losses	Quadratic	"Quadratic" Errors	3/2
Austerlitz	83	27	75	12	-16	-17	-16
Jena	74	4	43	12	3	2	5
Auershtedt	48	8	30	7	-3	-4	-3
Preisish	80	25	64	26	-6	-5	-2
Freiland	85	12	60	15	-2	-2	0
Aspern	75	25	70	35	6	7	8
Wagram	160	25	124	25	-6	-6	-3
Borodino	130	35	103	40	-6	-4	0
Berezina	75	6	45	15	2	3	5
Bautsen	163	18	96	12	-11	-11	-9
Ganau	75	15	50	9	-9	-9	-8
Drezden	160	20	125	15	-8	-9	-7
Keiptsig	300	50	200	60	-13	-10	-2
Katsbach	75	3	65	12	8	7	8
Liutsen	157	15	92	12	-8	-8	-6
Dennevits	70	9	57	9	-1	-2	-1
Kul'm	46	9	35	10	-1	-2	-1
Laon	100	2	45	6	1	0	2
Kpaon	30	18	18	5	-15	-15	-15
Waterloo	100	22	72	32	-2	1	5
Lun'i	120	11	85	11	-3	-4	-2
Grokhoro	72	9	56	12	0	0	1
Al'ma	62	3	34	6	1	0	1
Chernaia	62	2	56	8	6	5	5
Inkerman	90	12	63	3	-9	-10	-10
Col'ferino	170	20	150	18	-4	-5	-4
Madzhenta	58	10	54	5	-5	-6	-6
Kustotsa	70	8	51	8	-2	-3	-2

Battle	Stronger Force Start	Stronger Force Losses	Weaker Force Start	Weaker Force Losses	Quadratic	"Quadratic" Errors	3/2
Kenigrets	222	10	215	43	32	31	32
Mets	200	6	173	20	12	11	12
Gravelot	220	20	130	12	-13	-13	-11
Mars LaTour	125	16	65	16	-8	-8	-5
Vert	100	10	45	5	-7	-8	-7
Sedan	245	9	124	17	0	-1	3
Aladzha	60	2	36	15	6	7	9
Shabh	212	40	157	20	-25	-26	-23
Liaoian	150	18	120	24	1	1	3
Mukden	300	59	280	70	5	6	8

Quadratic	"Quadratic"	3/2
20.0%	15.0%	0.7%

Osipov's Errors

II.D. Fiske

Rear Admiral Bradley A. Fiske is regarded as a folk hero in the U.S. Navy. He was one of the primary operators in the military, technical, and political process of bringing the Navy into the Twentieth Century; of taking the technical advances of the late Nineteenth Century that made the transition from steam-driven wooden vessels to metal vessels possible. He was responsible for numerous naval inventions which spread the capabilities of modern technology through everyday maratine tasks. He was one of the architects of the operational innovations that integrated the new navy from a collection of ships into a viable military force. Thirdly, he was progenitor of the office of Chief of Naval Operations and the institutionalization of the General Staff in the U.S. Navy.

In 1905, Fiske wrote his eighty page essay "American Naval Policy" which was the Naval Institute (which he helped found, and of which he was later President,) prize essay of that year. In that essay, he introduced the concepts that we now think of as Lanchester's Quadratic Law (State Solution) and the Law of Concentration. This essay (until recently - see next section) gives rise to arguments that Fiske invented Attrition Theory.

While Fiske was prolific as an author, most of his writings have not been widely known outside of Naval circles. Of particular note, therefore, is the recent publication of Fiske's 1916/1918 **The Navy as a Fighting Machine.**[1] The 1916 edition met in 1917 enthusiastic review when published in England. From an European standpoint, this clearly makes Fiske a contemporary of Lanchester and Osipov in advancing (in print) the precepts of attrition theory.

In **The Navy as a Fighting Machine**, which incorporates an expansion of his prize essay as well as other material, Fiske discusses the implications of the Quadratic Law State Solution in a Naval context, much as Lanchester did with the Battle of Trafalgar, but in greater detail. He does not, however, extend his discussion to include any exact mathematical formalism of the state solution. (The 1918 edition notes the existence of such a formalism - see the next section.)

While he does not formulate an attrition theory in mathematical terms, Fiske does describe such a theory in words, and we can trans-late those words into mathematics. In particular, 'Fiske's attrition equations' take the form,

$$\begin{aligned} A(t+n\Delta t) &= A(t+(n-1)\Delta t) - \alpha\,\Delta t\; B(t+(n-1)\Delta t), \\ B(t+n\Delta t) &= B(t+(n-1)\Delta t) - \beta\,\Delta t\; A(t+(n-1)\Delta t), \end{aligned} \qquad \text{(II.D-1)}$$

where we have adopted the force strength and attrition rate notation (i.e. A, B and α, β) introduced earlier in describing Lanchester's attrition theory, and n here indicates the number of time periods of duration Δt which have transpired since battle began.

(Fiskian attrition is discrete in time rather than continuous as he describes it.)

Before proceeding with this discussion, it is useful at this time to define the finite difference operation. This is given by

$$\Delta A(t) = A(t+\Delta t) - A(t). \tag{II.D-2}$$

We will develop the finite difference formalism of attrition theory in a more complete manner in a later chapter.

By using the finite difference operator Δ, Equation (II.D-2), we may rewrite 'Fiske's attrition equations' as

$$\begin{aligned} \Delta A(t+(n-1)\Delta t) &= -\alpha\,\Delta t\; B(t+(n-1)\Delta t), \\ \Delta B(t+(n-1)\Delta t) &= -\beta\,\Delta t\; A(t+(n-1)\Delta t). \end{aligned} \tag{II.D-3}$$

We may 'read' these equations as: the change in the strength of a force (Fiske related this primarily to number of ships, but makes it clear that he is distinguishing combat power from mere numbers.) over a period of time Δt is negative (the force strength decreases,) and is equal to a damage coefficient (attrition rate constant/function multiplied by time period Δt - Lanchestrian terminology) times the strength of the opposing force at the beginning of the period. The right hand side of Equations (II.D-3) are the losses to the respective forces during the period.

(I have taken the liberty of introducing the damage coefficient to permit these equation to be written as equalities rather than as proportionalities as Fiske's discussion would literally indicate. He does discuss the damage causing process of combat and devotes considerable concern to the effectiveness of the units of the forces to cause damage - thus apparently not allowing the two forces to have distinct damage coefficients. For convenience of discussion, I have equated this damage coefficient to the attrition rate constant/ function multiplied by the time period. This allows the general form (for general $t' \equiv t + m\Delta t$) of Equations (II.D-3) to be rewritten (after a slight rearrangement,) as

$$\begin{aligned} \frac{\Delta A(t')}{\Delta t} &= -\alpha B(t'), \\ \frac{\Delta B(t')}{\Delta t} &= -\beta A(t'), \end{aligned} \tag{II.D-4}$$

which, if we take the limit as $\Delta t \to 0$ reduces to

$$\frac{dA(t')}{dt} = -\alpha B(t'),$$
$$\frac{dB(t')}{dt} = -\beta A(t'), \qquad \text{(II.D-5)}$$

since the left hand side of Equations (II.D-4) are, in the limit, just the definitions of the derivatives. Thus, Fiske's words are, in some approximate manner, mathematically equivalent to Lanchester's quadratic attrition differential equations, Equations (II.B-1) and (II.B-2).)

In his discussion of attrition, Fiske clearly identifies the condition that the 'damage coefficient' (attrition rate constant/ function) must truly be a constant. From a mathematical standpoint, this constraint of constancy represents an assumption for his analysis. Fiske further rightly notes that, for his analysis, knowledge of the length of the time period is unnecessary - rather, only the value of the 'damage coefficient' (he uses a value of 0.1 in his examples,) is necessary. This is correct only if supplemented by one more constraint - the 'damage coefficients' of the two forces are equal. Fiske explains this equality by citing the common armament (and thereby common damage caused by a hit,) of ships of the two forces. This view is reasonably well founded in terms of the historical development of warships in the period considered by Fiske.

Fiske notes that the duration of combat to a conclusion (in Lanchestrian terminology,) depends on the ratio of force strengths. He apparently arrived at this observation empirically from his examples rather than from analytical analyses such as would be possible from Osipov's explicit time solutions.

Finally, Fiske states that "**the difference in fighting forces cannot be measured in units of material and personnel only, though they furnish the most accurate general guide. Two other factors of great importance enter, the factors of skill and morale.**" In this regard, Fiske strikes the same note as Osipov.

Fiske also describes, in detail, what we know as the Principle of Concentration. He also states that "**every contest weakens the loser more that it does the winner**". This statement may be argued to be a corollary to Lanchester's Principle of Concentration and Osipov's Law of Distribution of Losses if we take the stronger force to be the likely winner (from an attrition sense.) In keeping with our previous discussions, we shall refer to this statement as Fiske's <u>Principle of Winning</u>.

While Fiske clearly has an earlier claim to the introduction of the concepts of Quadratic Law attrition theory, the scope of his contribution to the formalism of the theory is also clearly less than that of Lanchester and Osipov.

1. Fiske, Bradley A., Rear Admiral, U.S. Navy, **The Navy as a Fighting Machine**, Naval Institute Press, Annapolis, MD, 1988.

II.E. Chase

The fourth of our attrition theory pioneers is Jehu Valentine Chase. In a footnote[a] in his 1918 edition The Navy as a Fighting Machine, Fiske[1] cites (then) Lt. Chase's 1902 Naval War College Paper "Sea Fights: A Mathematical Investigation of the Effect of Superiority of Force in". This brief mathematical paper (|| 3 pages) was initially classified and was not declassified until 1972. Wayne Hughes (CAPT., U.S. Navy Ret.), one of the editors of the 1988 republication of Fiske's book, includes this essay as an appendix[2] and decries the hiding of this work. Surely, in light of the publication of Lanchester's book, this continued safeguarding of Fiske's document for those 56 years must come under question.

Also included in the appendix is an extract from a 1921 letter written by (then) CAPT. Chase (He eventually held the rank of Rear Admiral.) in which he discusses the Quadratic Law/Principle of Concentration and the counteracting considerations of survivability of the force in terms of how a Naval force is designed - many smaller ships are more survivable than a few small ships. (The question of survivability in the context of attrition theory is a subject which we will treat in a later chapter.)

In his original paper, Chase describes his own version of Quadratic Law attrition. To do this, he first introduces the concept of "sudden" versus "continuous gradual destruction" (i.e. attrition). In modern terminology, continuous gradual destruction is usually referred to as "graceful degradation".[3] In brief, this concept holds that the attrition of units (or more generally, reduction of system performance,) occurs in an essentially continual manner. The concept of sudden destruction holds that attrition is punctuated and total - a unit is either totally effective or totally ineffective, and the change occurs over a short period of time (often treated as instantaneous.) An example of this which is frequently offered is the attrition of tanks by modern weapons. Until a tank is hit, its effectiveness is not usually considered to be diminished; however, once the tank is hit, the probability of kill given a hit is sufficiently great (in most cases,) that the tank is "killed". This occurs over a period of time which is of the order of fractions of a second. (The consideration of the transition from sudden to continuous gradual attrition is a subject of great importance in the conjugate theory of attrition rate constants/functions.)

If a unit, on the other hand, is comprised of several tanks, then the unit is not "killed" until all the tanks in the unit have been individually "killed". Further, each time that a tank is "killed", the effectiveness of the unit is reduced by an amount approximately equal to the fraction of the unit that the tank represents (for a ten tank company. each "kill" reduces the effectiveness of the unit by 10% - this view

[a] Isn't it amazing that the most interesting information comes from footnotes? Both Osipov and Chase were introduced that way.

neglects any contribution to the unit's effectiveness of morale or other psychological, training, or tactical influences.) Nonetheless, this simple example illustrates the basic idea of continuous gradual destruction. Such a concept is applicable to Naval warships which have a large number of weapon systems and other assets such as engines, ammunition stores, and command and control systems which contribute to its total effectiveness and which effectiveness is only completely exhausted when some sizable portion of the ship's weapon systems and other assets are rendered individually ineffective.

Chase acknowledges that sudden destruction does occur for ships (which he was solely concerned with) due to ramming, running aground at speed, or torpedo impact (for smaller ships,) but that for gunfire, attrition of the ship as a whole is gradual. In other words, it takes several (many) gunfire hits to disable a ship. Since these hits may be presumed to impact in a random manner, [We will consider the statistics of this process in a later chapter on attrition processes.] the actually punctuated but drawn out process can be approximated as a continuous process.

Chase defines the following quantities:

m,n are the number of ships on each side (that are engaged in combat with each other,)

a_m, a_n are the units of "life" of each type of ship (each side is implicitly assumed to have only one type of ship, but the two sides may each be comprised of a different type of ship - this reflects the continued, at that time, theory of using the Line of Battle and the fact that ships are usually produced in series with relatively little difference among ships of the same series,)

b_m, b_n are the units of "destruction" per time which each ship (of each side) can produce,

D_m, D_n are the damage received by each ship (at a given instant of time,) and

y, z are the "destructive power" of each m, n ship at a given instant of time.

[I have taken the liberty of changing the subscripts designating the two forces from the numbers used by Chase to letters to reduce confusion.]

In addition, total damage is spread equally over all ships on a given side; ships are tacitly assumed never to actually sink (this is a moot point and open to some interpretation,) the units of "destruction" may be thought of as essentially the number of 'independent' [We will define this distinction in a later chapter, however, repeated hits on an already destroyed weapon system or asset cause little additional damage and are thus not "independent" in reducing the effectiveness of the ship.] hits per time, and the units of "life" as the number of hits that a ship may take before it can no longer fight (sink?)

Chase provides the relational equations

$$a_m\, y = b_m\,(a_m - D_m),$$
$$a_n\, z = b_n\,(a_n - D_n), \tag{II.E-1}$$

which state that the product of the number of "life" units and the instantaneous "destructive power" of a ship are equal to the product of the "destructive" rate of that ship and the difference between the "life" of the ship and the damage the ship has received. In words, this equation is

(initial "life")(instantaneous "destructive power)
= ("destructive" rate)(instantaneous "life" remaining)

If we note that the damage received D_m, D_n, and the "destructive power" y, z, of each ship are time dependent, we may use the definitions of the "destructive power",

$$D_m(t) = \frac{m}{n}\int_0^t z(t')\,dt',$$
$$D_n(t) = \frac{n}{m}\int_0^t y(t')\,dt', \tag{II.E-2}$$

to form pairs of "attrition" differential equations in D_m, D_n, or y, z. (We will not treat these differential equations explicitly here since they were not part of Chase's exposition, but delay their explicit solution for a later chapter.)

Chase then equates Equations (II.E-1) and (II.E-2) (appropriately,) and differentiates with respect to time. This gives

$$\frac{a_m}{b_m}\frac{dy}{dt} = -\frac{n}{m}\,z,$$
$$\frac{a_n}{b_n}\frac{dz}{dt} = -\frac{m}{n}\,y, \tag{II.E-3}$$

which are Quadratic Law-type attrition differential equations. Time may be removed from these equations to yield the single differential equation,

$$\frac{a_m\, b_n}{a_n\, b_m}\frac{dy}{dz} = \frac{n^2}{m^2}\frac{z}{y}, \tag{II.E-4}$$

which may be written in an exact form,

$$\frac{a_m\, b_n}{a_n\, b_m}\, y\, dy = \frac{n^2}{m^2}\, z\, dz. \tag{II.E-5}$$

Rather than integrate this in the usual definite form, Chase does the integration indefinitely to yield

$$\frac{a_m\, b_n}{a_n\, b_m}\, y^2 = \frac{n^2}{m^2}\, z^2 + C, \tag{II.E-6}$$

which is the state solution for Chase attrition. The boundary conditions on Equation (II.E-6) may be found by examining Equations (II.E-1) and (II.E-2), and noting that at $t = 0$, D_m and D_n are zero, (assuming y and z are well defined and behaved in a mathematical sense.) Thus, at $t = 0$, $y = b_m$ and $z = b_n$. This gives a value for C of

$$C = b_n^2 \left[\frac{a_m\, b_m}{a_n\, b_n} - \frac{n^2}{m^2}\right]. \tag{II.E-7}$$

If the battle continues to a conclusion (in a Lanchestrian sense - the concept is independently introduced by Chase without comment,) then the "destructive power" of one side becomes zero. Chase selects $z = 0$ at the conclusion; this gives a state solution

$$\frac{a_m\, b_n}{a_n\, b_m}\, y^2 = b_n^2 \left[\frac{a_n\, b_n}{a_m\, b_m} - \frac{n^2}{m^2}\right]. \tag{II.E-8}$$

Chase then uses this equation to solve for the damage received by each ship of the surviving force at the conclusion,

$$D_m = \frac{a_m - \sqrt{\dfrac{a_m}{b_m}\,(m^2 a_m b_m - n^2 a_n b_n)}}{m}. \tag{II.E-9}$$

Since the damage received by each ship of the destroyed force is just

$$D_n = a_n, \tag{II.E-10}$$

by implication of the conclusion condition (total destruction!) the ratio of the total damage to the surviving force to the total damage to the destroyed force is

$$\frac{m\, D_m}{n\, D_n} = \frac{a_m}{a_n}\left[\frac{m - \sqrt{m^2 - n^2 \dfrac{a_n b_n}{a_m b_m}}}{n}\right]. \tag{II.E-11}$$

Chase also considers the case of a draw (where the two fleets are equally matched.) This gives, from Equation (II.E-9) (since y and z are both zero at

conclusion,)

$$m^2 \, a_m \, b_m = n^2 \, a_n \, b_n, \tag{II.E-12}$$

Finally, if the ships on both sides are equivalent (Chase's term is "similar" - life and damage rates are equal for the two forces,) the total damage ratio becomes

$$\frac{m \, D_m}{n \, D_n} = \frac{m - \sqrt{m^2 - n^2}}{n}, \tag{II.E-13}$$

while the draw condition becomes

$$n = m. \tag{II.E-14}$$

It is illuminating that Chase does not elaborate his mathematics with explanation - apparently he felt such to be unnecessary. As such, he represents the opposite extreme from the other three pioneers, especially Fiske.

1. Fiske, Bradley A., Rear Admiral, U.S. Navy, **The Navy as a Fighting Machine**, Naval Institute Press, Annapolis, MD, 1988.

2. "Lieutenant J. V. Chase's Force-on-Force Effectiveness Model for Battle Lines", Appendix C in Fiske.

3. Callahan, Leslie G., Jr., Ph.D., and COL (USA Ret.), "Modeling, Simulation and Gaming of Warfare - Course Overview", Ninth Annual Course on Modeling, Simulation and Gaming of Warfare, Georgia Institute of Technology, Atlanta, GA, August 1988.

II.F. Conclusion

Of the three pioneers, Chase clearly has the claim for earliest advancement of attrition theory. The classification of his paper, removing it from public consideration compromises that claim, effectively reducing that claim to an academic footnote. Had his paper not been hidden, the terseness of the development would have limited it to a military audience with mathematical faculty and intellectual inquisitiveness adequate to flesh out the theory - a markedly more limited community than that which could read and debate the works of the other three pioneers. Albeit, an argument may be raised that had the paper not been classified, Chase would have expanded his brief paper into a more robust exposition of attrition theory. As fetching as this argument may be, especially in terms of its effect on subsequent history, such considerations are of the nature of science fiction, and the fact remains that Chase's work was buried from the light of scientific day.

Neglecting therefore, Chase's claim to primacy, the question still remains of which pioneer should be considered to be first? If we compare the works of the other three, there is still Fiske's 1905 prize essay which first introduced the Quadratic Law concept but lacked an firm mathematical underpinning (nor did the 1916/1918 book rectify this shortfall.) Next appears to be Lanchester with his 1914 article, followed by Osipov with his series of articles in 1915. Both Lanchester and Osipov clearly laid down firm mathematical bases for their theories. Both clearly built different frameworks around their theories.

The question of primacy is moot and cloudy. Chase published first and had the claim of primacy effectively denied him by government instrumentality. Fiske clearly published second but failed to provide a mathematical formalism. Lanchester and Osipov published next, within months of each other. Concurrency of their work cannot be easily dismissed from what we know today. If we consider the impact of the publications on the public, it is clear that Fiske and Lanchester (based on Chase's letter of 1921,) were the better known. Chase was known only in cleared U.S. Navy circles and Osipov was known only in Russia (?). Thus, we come full circle, finding that the best claim to being 'father' of attrition theory seems to be Lanchester's.

None of this discussion of primacy is really meaningful. Who was first is not really a measure of who (or what) is important to attrition theory. Regardless of who we select as 'father', and for traditional reasons, we will continue to use Lanchester and the permutations and labels based on his name as the standard, what is really important are the contributions of these pioneers to the theory of attrition. These contributions are considerable, including the mathematical theories of Chase, Lanchester, and Osipov, Lanchester's Principle of Concentration, Osipov's Law of Distribution of Losses, Fiske's Principle of Winning, and Chase's Concept of Continuous Gradual Destruction. These and other contributions, and developments from these are the subject matter of the rest of this book.

III. MATHEMATICAL THEORY I:
Fundamental Solutions of the Lanchester Attrition Differential Equations

III.A. Introduction

In this section, we present a brief review of the mathematical methods used in solving the Lanchester differential equations as they have been presented thus far.

As stated in the previous chapter, the general form of the Lanchester differential equations is

$$\frac{dA}{dt} = -\alpha A^{2-n} B, \qquad \frac{dB}{dt} = -\beta B^{2-n} A. \tag{III.A-1}$$

As part of this, we will be concerned with three pairs of differential equations which give rise to: the linear law

$$\frac{dA}{dt} = -\alpha A B, \tag{III.A-2}$$

and

$$\frac{dB}{dt} = -\beta B A; \tag{III.A-3}$$

the square (quadratic) law

$$\frac{dA}{dt} = -\alpha B, \tag{III.A-4}$$

and

$$\frac{dB}{dt} = -\beta A; \tag{III.A-5}$$

and the mixed law

$$\frac{dA}{dt} = -\alpha B, \tag{III.A-6}$$

and

$$\frac{dB}{dt} = -\beta AB. \qquad \text{(III.A-7)}$$

The first two sets of these differential equations, which give rise to the linear and square laws, have an exchange symmetry of the form $(\alpha, B_0) \Leftrightarrow (\beta, A_0)$ which allow the construction of the mathematical form of the second solution, B(t), from the mathematical form of the first solution, A(t), by the use of this symmetry. The differential equations giving rise to the mixed law do not possess such a symmetry and the mathematical forms of the two solutions, A(t) and B(t), must be constructed separately.

III.B. State Solutions

If the force strengths are assumed to be explicit functions of time, then each pair of differential equations above may be combined into one equation by removing time as a variable. This, in the linear law case, we may write

$$\begin{aligned}\frac{\frac{dA}{dt}}{\frac{dB}{dt}} &= \frac{dA}{dt}\,\frac{dt}{dB} \\ &= \frac{dA}{dB} = \frac{-\alpha AB}{-\beta BA} = \frac{\alpha AB}{\beta AB} \\ &= \frac{\alpha}{\beta}.\end{aligned} \qquad \text{(III.B-1)}$$

This equation may be integrated directly as

$$\beta \int_{A_0}^{A} dA' = \alpha \int_{B_0}^{B} dB', \qquad \text{(III.B-2)}$$

which yields

$$\beta(A - A_0) = \alpha(B - B_0), \qquad \text{(III.B-3)}$$

from which the origin of the term 'linear law' may be clearly seen since this is the equation of a straight line. Equation (III.B-3) is known as the state solution for the linear law.

Equation (III.B-3) simply states that the strength of one force (say A) at any

time t is linearly related to the strength of the other force at the same time. This equation tells us very little (per se) of the time dependence of A or B, only about their mutual and direct dependence on each other.

The 'square law' differential equations may be solved in the same manner for the differential equations (III.A-4) and (III.A-5):

$$\frac{\frac{dA}{dt}}{\frac{dB}{dt}} = \frac{dA}{dt}\frac{dt}{dB} = \frac{dA}{dB} = \frac{-\alpha B}{-\beta A} = \frac{\alpha B}{\beta A}. \tag{III.B-4}$$

This equation may be integrated directly as

$$\beta \int_{A_0}^{A} A'\, dA' = \alpha \int_{B_0}^{B} B'\, dB', \tag{III.B-5}$$

which yields

$$\beta (A^2 - A_0^2) = \alpha (B^2 - B_0^2), \tag{III.B-6}$$

which is the 'square law' state solution. (Normally, the factors of 2 in the denominators are dropped since they occur on both sides of the equation.)

The 'mixed law' differential equations may be solved in the same manner as the previous two, by elimination of parametric time:

$$\frac{\frac{dA}{dt}}{\frac{dB}{dt}} = \frac{dA}{dt}\frac{dt}{dB} = \frac{dA}{dB} = \frac{-\alpha B}{-\beta BA} = \frac{\alpha}{\beta A}. \tag{III.B-7}$$

This equation may be integrated directly as

$$\beta \int_{A_0}^{A} A' \, dA' = \alpha \int_{B_0}^{B} dB', \tag{III.B-8}$$

which yields

$$\frac{\beta}{2}(A^2 - A_0^2) = \alpha (B - B_0), \tag{III.B-9}$$

which demonstrates the 'mixed' nature of the state solution of differential equations (III.A-6) and (III.A-7).

Several mathematical insights may be drawn from the state solutions. One of the most common of these is the development of so-called victory conditions. If it is assumed that the two forces battle until only one force remains, and that complete annihilation (battle to a conclusion in a Lanchestrian sense,) may be called victory, then equations (III.B-3), (III.B-6) and (III.B-9) may be rewritten in the forms:

$$\alpha B_0 - \beta A_0 = \alpha B - \beta A, \tag{III.B-10}$$

$$\alpha B_0^2 - \beta A_0^2 = \alpha B^2 - \beta A^2, \tag{III.B-11}$$

and

$$\alpha B_0 - \frac{\beta}{2} A_0^2 = \alpha B - \frac{\beta}{2} A^2, \tag{III.B-12}$$

where all of the initial force strengths ('$_0$' subscripted terms) have been moved onto the left hand side of equations (III.B-10) - (III.B-12). Since the two forces battle until only one retains any strength, then either A = 0, or B = 0, at the battle's end. Thus, in all three cases, the right hand side of these equations are either positive or negative depending on whether B or A 'wins' (respectively.) That is, if B 'wins', the right hand side of any of these three equations will be positive, while if A 'wins', the right hand side of any of the equations will be negative by virtue of the minus sign. (Recall that A and B are by definition nonnegative.)

Notice that since these are 'equations', the same conditions must apply to the left hand side as to the right. We may thus write:

$$\alpha B_0 - \beta A_0 > 0 \quad (B\ wins) \qquad \text{(III.B-13)}$$
$$< 0 \quad (A\ wins),$$

$$\alpha B_0^2 - \beta A_0^2 > 0 \quad (B\ wins) \qquad \text{(III.B-14)}$$
$$< 0 \quad (A\ wins),$$

and

$$\alpha B_0 - \frac{\beta}{2} A_0^2 > 0 \quad (B\ wins) \qquad \text{(III.B-15)}$$
$$< 0 \quad (A\ wins).$$

By rearrangement then, we can write

$$\frac{\alpha B_0}{\beta A_0} > 1 \quad (B\ wins) \qquad \text{(III.B-16)}$$
$$< 1 \quad (A\ wins),$$

$$\frac{\alpha B_0^2}{\beta A_0^2} > 1 \quad (B\ wins) \qquad \text{(III.B-17)}$$
$$< 1 \quad (A\ wins),$$

and

$$\frac{2\alpha B_0}{\beta A_0^2} > 1 \quad (B\ wins) \qquad \text{(III.B-18)}$$
$$< 1 \quad (A\ wins),$$

If any of these three (equations (III.B-16) - (III.B-18)) are equalities, then the prediction is for a draw or 'tie' (mutual annihilation?)

It may be noted that these 'victory' conditions are for a battle where one force is completely destroyed. From an historical standpoint, such battles are relatively rare. We shall examine conclusive battles in a later chapter.

Further, as we shall discuss later, even when one force is reduced completely, the form of the relevant differential equations seem to be changed.

III.C. Direct Methods of Solution

The state solutions are useful, but they convey imperfect information about the actual time dependence of the force strengths. In this section, we shall address two direct methods of solution of the square law differential equations, the method of substitution and the method of Frobenius.

Before continuing, it is worthwhile noting that the Lanchester differential equations are first order only. This means that only one boundary condition may be imposed on each solution. As we shall see, this sometimes leads to some less than satisfying conditions. It does have the satisfying result of assuring us that the solution we arrive at is the unique solution.

III.C.1. Method of Substitution

The square law differential equations (Equations (III.A-4) and (III.A-5)) may be solved directly by substitution. If we take one of the two differential equations and differentiate it with respect to time, we get

$$\frac{d^2A}{dt^2} = -\frac{d\alpha}{dt}B - \alpha\frac{dB}{dt}. \qquad \text{(III.C-1)}$$

We normally assume that the attrition coefficients are time independent,[a] so the first term on the right hand side of Equation (III.C-1) is zero. The second right hand side term is just the other Lanchester differential equation of the pair. If we substitute Equation (III.A-5) into this equation, we get

$$\frac{d^2A}{dt^2} = -\alpha\beta A, \qquad \text{(III.C-2)}$$

and if we define:

$$\alpha\beta \equiv \gamma^2, \qquad \text{(III.C-3)}$$

we see that the resulting differential equation (of the second order)

---

[a] We shall consider time (and range,) dependent attrition coefficients in later chapters.

$$\frac{d^2A}{dt^2} = \gamma^2 A, \tag{III.C-4}$$

has the solutions

$$A(t) = Ce^{\gamma t} + De^{-\gamma t}, \tag{III.C-5}$$

Since we will be applying initial conditions, (making A and B take on values of A_0 and B_0 at t = 0,) it is more useful to write the solution as

$$A(t) = C \sinh(\gamma t) + D \cosh(\gamma t) \tag{III.C-6}$$

and we can calculate the solution for B by either direct differentiation of Equation (III.C-6), or by symmetry. If we calculate it by differentiation, the solution may be immediately seen as

$$B(t) = -\frac{\gamma}{\alpha} D \sinh(\gamma t) - \frac{\gamma}{\alpha} C \cosh(\gamma t). \tag{III.C-7}$$

(Note that even though we have a second order differential equation, the boundary (in this case, initial) conditions imposed are the same as would be imposed for the pair of first order differential equations. Thus, we are neither requiring nor introducing any new information. That is, we require Equations (III.C-6) (A(t)) and (III.C-7) (the derivative of A(t) or just B(t)) to have the proper behavior at t = 0.)

We now invoke the properties of the hyperbolic sine and cosine, namely that
sinh(0) = 0, and
cosh(0) = 1,
to write:

$$A_0 = D, \tag{III.C-8}$$

and

$$B_0 = -\frac{\gamma}{\alpha} C, \tag{III.C-9}$$

which may be substituted back into Equations (III.C-6) and (III.C-7) to yield:

$$A(t) = A_0 \cosh(\gamma t) - \sqrt{\frac{\alpha}{\beta}}\, B_0 \sinh(\gamma t), \tag{III.C-10}$$

and

$$B(t) = B_0 \cosh(\gamma t) - \sqrt{\frac{\beta}{\alpha}}\, A_0 \sinh(\gamma t), \tag{III.C-11}$$

from which the symmetry of exchange of $(\alpha, B_0) \Leftrightarrow (\beta, A_0)$ is obvious. These two equations are the explicit time solutions of the Lanchester square law differential equations.

The linear and mixed law differential equations cannot be solved by substitution since they continually mix the two force strengths with repeated differentiation. Thus, the method of substitution is of value only in solving the linear law differential equations.

III.C.2. Method of Frobenius

In the solution by the method of Frobenius, we assume that the time solutions of the Lanchester differential equations may be represented as power series in time:

$$A(t) = \sum_{n=0}^{\infty} a_n\, t^n, \tag{III.C-12}$$

and similarly for B(t).

If we differentiate the series and substitute them into the square law differential equations, Equations (III.A-4) and (III.A-5), we get (after adjusting the indices on the left,)

$$\sum_{n=0}^{\infty} (n+1)\, a_{n+1}\, t^n = -\alpha \sum_{n=0}^{\infty} b_n\, t^n, \tag{III.C-13}$$

and

$$\sum_{n=0}^{\infty} (n+1)\, b_{n+1}\, t^n = -\beta \sum_{n=0}^{\infty} a_n\, t^n. \tag{III.C-14}$$

If we now require that each term in the series be linearly independent, we may equate terms with common powers of t. This gives,

$$a_{n+1} = -\frac{\alpha}{n+1}\, b_n, \tag{III.C-15}$$

and

$$b_{n+1} = -\frac{\beta}{n+1}\, a_n. \tag{III.C-16}$$

By adjusting indices and substituting, this becomes

$$a_{n+2} = \frac{\gamma^2}{(n+1)(n+2)}\, a_n, \tag{III.C-17}$$

and similarly for the b_n.

From the initial conditions,

$a_0 = A_0,$
$b_0 = B_0,$

and

$a_1 = -\alpha B_0,$
$b_1 = -\beta A_0.$

We notice immediately that the result is an alternating series in odd and even powers of t. That is:

$$\begin{aligned} a_2 &= \frac{\gamma^2}{(2)(1)} A_0, \\ a_3 &= -\alpha\, \frac{\gamma^2}{(3)(2)}\, B_0, \\ a_4 &= \frac{\gamma^4}{(4)(3)(2)(1)}\, A_0, \\ a_5 &= -\alpha\, \frac{\gamma^4}{(5)(4)(3)(2)}\, B_0, \end{aligned} \tag{III.C-18}$$

or

$$\begin{aligned} a_n &= \frac{\gamma^n}{n!} A_0, \quad (n\ even) \\ &= -\alpha\, \frac{\gamma^{n-1}}{n!}\, B_0, \quad (n\ odd) \\ &= -\sqrt{\frac{\alpha}{\beta}}\, \frac{\gamma^n}{n!} B_0, \quad (n\ odd) \end{aligned} \tag{III.C-19}$$

and upon substitution back into the series, this yields

$$A(t) = A_0 \sum_{\substack{n=0 \\ n\ \text{even}}}^{\infty} \frac{\gamma^n t^n}{n!} - \sqrt{\frac{\alpha}{\beta}}\, B_0 \sum_{\substack{n=0 \\ n\ \text{odd}}}^{\infty} \frac{\gamma^n t^n}{n!} \tag{III.C-20}$$

$$= A_0 \cosh(\gamma t) - \sqrt{\frac{\alpha}{\beta}}\, B_0 \sinh(\gamma t),$$

which is the same as Equation (III.C-11).

Unfortunately, the method of Frobenius is also not useful for solving the linear and mixed equations because the differential equations are not linear. To solve these differential equations, we must turn to some other, more general method to find solutions to the other Lanchester differential equations.

III.D. Normal Forms

The method that permits general solution of the Lanchester differential equations presented thus far is the normal forms method. It is so called because the state solutions of the differential equations must be developed first.

Before proceeding, we note that for the square and linear laws, an exchange symmetry $(\alpha, B_0) \Leftrightarrow (\beta, A_0)$ exists. Because of this symmetry, we shall not have to explicitly derive solutions for both of the differential equations of these pairs. This symmetry, sadly, is not the case for the mixed law, and solutions for both of these differential equations will have to be developed.

III.D.1 Linear Lanchester Equations

To demonstrate the normal forms method of solution, we begin with one of the linear law differential equations (Equations (III.A-2) and (III.A-3)) and write the direct integration solution as

$$\int_{A_0}^{A} \frac{dA'}{\alpha\, A'\, B'} = -\int_{t_0}^{t} dt', \tag{III.D-1}$$

and we rewrite Equation (III.B-3), the state solution as

$$\alpha B' = \beta A' - \beta A_0 + \alpha B_0, \tag{III.D-2}$$

and substitute the state solution directly into the denominator of the left hand side of

Equation (III.D-1) to yield

$$\int_{A_0}^{A} \frac{dA'}{\beta A' - \beta A_0 + \alpha B_0} = -\int_{t_0}^{t} dt', \tag{III.D-3}$$

the right hand side integral of this equation can be performed directly as an elementary integral; the left hand side integral may be taken from Appendix A , integral (A-1) with parameters:

a = β

b = βA_0 - αB_0 $\equiv$ Δ_1 (a 'victory' condition (conclusion) statement).

The resulting integrations have the form

$$-\frac{2}{\Delta_1} \coth^{-1}\left(\frac{2\beta A}{\Delta_1} - 1\right)\Bigg|_{A_0}^{A(t)} = -\Delta t, \tag{III.D-4}$$

where : $\Delta t = t - t_0$.

Substitution of the limits on the right hand side and rearrangement yield

$$\coth^{-1}\left(\frac{2\beta A(t)}{\Delta_1} - 1\right) = \coth^{-1}\left(\frac{2\beta A_0}{\Delta_1} - 1\right) + \frac{\Delta_1}{2}\Delta t, \tag{III.D-5}$$

We may now make use of the identity

$$\coth^{-1}(x) = \frac{1}{2} \ln\left(\frac{x+1}{x-1}\right), \tag{III.D-6}$$

to write (after some rearrangement)

$$A(t) = A_0 \frac{\Delta_1}{\beta A_0 - \alpha B_0 e^{-\Delta_1 \Delta t}}. \tag{III.D-7}$$

The solution for B(t) can be formed from Equation (III.D-7) by using the symmetry properties; that is, by swapping α and β, and A_0 and B_0.

III.D.2 Square Lanchester Equations

The square law differential equations may be solved in the same manner. We may rewrite Equation (III.B-6) as

$$B' = \sqrt{\frac{\beta}{\alpha}A'^2 + B_0^2 - \frac{\beta}{\alpha}A_0^2}. \tag{III.D-8}$$

and substitute it into a direct integration solution of Equation (III.A-4). This yields

$$\int_{A_0}^{A(t)} \frac{dA'}{\sqrt{\frac{\beta}{\alpha}A'^2 - \frac{\beta}{\alpha}A_0^2 + B_0^2}} = -\alpha\int_{t_0}^{t} dt'. \tag{III.D-9}$$

The left hand side integral is again found in Appendix A, integral (A-2) with parameters

$$a = \frac{\beta}{\alpha}, \qquad b = B_0^2 - \frac{\beta}{\alpha}A_0^2, \tag{III.D-10}$$

and define:

$$\Delta_2 = \alpha B_0^2 - \beta A_0^2. \tag{III.D-11}$$

Evaluation of the integrals yields

$$\sqrt{\frac{\beta}{\alpha}}\ \sinh^{-1}\left(\sqrt{\frac{\beta}{\Delta_2}}\,A'\right)\Bigg|_{A_0}^{A(t)} = -\alpha\,\Delta t, \tag{III.D-12}$$

which rearranges to

$$\sinh^{-1}\left(\sqrt{\frac{\beta}{\Delta_2}}\,A(t)\right) = \sinh^{-1}\left(\sqrt{\frac{\beta}{\Delta_2}}\,A_0\right) - \gamma\,\Delta t, \tag{III.D-13}$$

and use the identity

$$\sinh(u - v) = \sinh(u)\cosh(v) - \cosh(u)\sinh(u), \tag{III.D-14}$$

to get

$$\sqrt{\frac{\beta}{\Delta_2}}\,A(t) = \sqrt{\frac{\beta}{\Delta_1}}\,A_0 \cosh(\gamma\,\Delta t) - \cosh\left[\sinh^{-1}\left(\sqrt{\frac{\beta}{\Delta_1}}\,A_0\right)\right] \sinh(\gamma\,\Delta t), \tag{III.D-15}$$

and use the identity

$$\sinh^{-1}(x) = \cosh^{-1}\left(\sqrt{x^2+1}\right), \tag{III.D-16}$$

and the state solution to reduce this to

$$A(t) = A_0 \cosh(\gamma\,\Delta t) - \sqrt{\frac{\alpha}{\beta}}\,B_0 \sinh(\gamma\,\Delta t). \tag{III.D-17}$$

identical to Equation (III.C-10), the time solution of the square law Lanchester differential equations.

III.D.3 Mixed Lanchester Equations

The final exercise of the indirect method is the solution of the mixed law Lanchester differential equations, Equations (III.A-6) and (III.A-7). To arrive at this solution, we must rewrite the state solutions of the mixed law, Equation (III.B-9) as,

$$B' = \frac{\beta}{2\alpha}\,A'^2 + \frac{\Delta_m}{\alpha}, \tag{III.D-18}$$

where:

$$\Delta_m \equiv \alpha B_0 - \frac{\beta}{2}A_0^2. \tag{III.D-19}$$

We substitute Equation (III.D-18) into the direct solution

$$\int_{A_0}^{A(t)} \frac{dA'}{\frac{\beta}{2}A'^2 + \frac{\Delta_m}{\alpha}} = -\int_{t_0}^{t} dt'. \tag{III.D-20}$$

This integral has two different forms depending on whether Δ_m is greater or less than zero. We shall treat the former case first. If we make use of integral (A-3) from

Appendix A and apply the argument addition rule for tangents, we may write

$$A(t) = \frac{A_0 - \sqrt{\dfrac{2|\Delta_m|}{\beta}}\tan(\eta t)}{1 + A_0\sqrt{\dfrac{\beta}{2|\Delta_m|}}\tan(\eta t)}, \ \Delta_m > 0, \tag{III.D-21}$$

where:

$$\eta = \sqrt{\frac{\beta|\Delta_m|}{2}}. \tag{III.D-22}$$

The A(t) solution when $\Delta_m < 0$, may be derived from Equation (III.D-21) by noting that when Δ_m becomes negative, then $\eta \to i\eta$, and that $\tan(i\eta t) = i\tanh(\eta t)$. Thus, we may write

$$A(t) = \frac{A_0 + \sqrt{\dfrac{2|\Delta_m|}{\beta}}\tanh(\eta t)}{1 + A_0\sqrt{\dfrac{\beta}{2|\Delta_m|}}\tanh(\eta t)}, \ \Delta_m < 0. \tag{III.D-23}$$

The B(t) solution can be found by either performing the normal form integration of the other attrition differential equation, or by substituting Equations (III.D-21) and (III.D-23), respectively, back into the rewritten attrition differential equation,

$$B(t) = -\frac{1}{\alpha}\frac{dA}{dt}. \tag{III.D-24}$$

This allows us to write the two solutions, after some algebra, as

$$B(t) = \frac{B_0\sec^2(\eta t)}{\left[1 + A_0\sqrt{\dfrac{\beta}{2|\Delta_m|}}\tan(\eta t)\right]^2}, \ \Delta_m > 0, \tag{III.D-25}$$

and

$$B(t) = \frac{B_0 \, sech^2(\eta t)}{\left[1 + A_0 \sqrt{\frac{\beta}{2|\Delta_m|}} \tanh(\eta t)\right]^2}, \; \Delta_m < 0, \qquad \text{(III.D-26)}$$

III.E. Force Ratio

One of the quantities which is of interest in attrition theory is the force ratio; that is, the ratio of the two force strengths. If the force ratio is represented by ρ(t), then it is defined by

$$\rho(t) = \frac{A(t)}{B(t)}. \qquad \text{(III.E-1)}$$

Before calculating the derivative of this quantity to form its attrition differential equation, we note in passing that, at most, the quadratic Lanchester differential equations will possess a closed form solution for the force ratio, but not either the linear or the mixed Lanchester differential equations. This sad situation is predicted by the fact that the time dependent solutions of both of these sets of differential equations contain their state solutions explicitly in the time dependent portions of the solutions. Only the quadratic solutions do not contain the state solution in such a way.

We may calculate the time derivative of the force ratio,

$$\frac{d\rho}{dt} = \frac{1}{B}\frac{dA}{dt} - \frac{A}{B^2}\frac{dB}{dt}, \qquad \text{(III.E-2)}$$

into which we may substitute Equations (III.A-1) to yield,

$$\frac{d\rho}{dt} = -\alpha A^{2-n} + \beta B^{-n} A^2, \qquad \text{(III.E-3)}$$

from which we may see that the right hand side reduces to a function of ρ only if n = 2! Thus,

$$\frac{d\rho}{dt} = \beta \rho^2 - \alpha. \qquad \text{(III.E-4)}$$

We may solve this exact differential equation using the same techniques that we used for the mixed Lanchester differential equations, giving a solution

$$\rho(t) = \frac{\rho_0 - \delta \tanh(\gamma t)}{1 - \dfrac{\rho_0 \tanh(\gamma t)}{\delta}}. \qquad \text{(III.E-5)}$$

This result could, of course, have been derived directly from Equations (III.C-10) and (III.C-11), although that method would not have been as theoretically useful. The force ratios of the other two types of attrition, linear and mixed, can be formed by direct ratioing; however, the resulting ratios cannot be mathematically manipulated to remove the initial force strengths explicitly.

III.F. Summary of Solutions

This concludes the development of the basic solutions of the Lanchester differential equations. We present here, for the use of those who do not choose to follow the mathematics, or who may wish to use these as a reference, a summary of the relevant solutions for each set of differential equations:

Linear Equations

$$A(t) = A_0 \frac{\Delta_1}{\beta A_0 - \alpha B_0 e^{-\Delta_1 \Delta t}}. \qquad \text{(III.F-1)}$$

$$B(t) = B_0 \frac{-\Delta_1}{\alpha B_0 - \beta A_0 e^{\Delta_1 \Delta t}}. \qquad \text{(III.F-2)}$$

$$\Delta_1 \equiv \beta A_0 - \alpha B_0 \qquad \text{(III.F-3)}$$

Quadratic Equations

$$A(t) = A_0 \cosh(\gamma t) - \delta\, B_0 \sinh(\gamma t), \qquad \text{(III.F-4)}$$

$$B(t) = B_0 \cosh(\gamma t) - \frac{1}{\delta} A_0 \sinh(\gamma t), \qquad \text{(III.F-5)}$$

$$\rho(t) = \frac{\rho_0 - \delta \tanh(\gamma t)}{1 - \dfrac{\rho_0 \tanh(\gamma t)}{\delta}}. \qquad \text{(III.F-6)}$$

$$\Delta_2 \equiv \alpha B_0^2 - \beta A_0^2 \qquad \text{(III.F-7)}$$

$$\gamma \equiv \sqrt{\alpha \beta} \qquad \text{(III.F-8)}$$

$$\delta \equiv \sqrt{\frac{\alpha}{\beta}} \qquad \text{(III.F-9)}$$

$$\rho(t) = \frac{A(t)}{B(t)} \qquad \text{(III.F-10)}$$

Mixed Equations

$$A(t) = \frac{A_0 - \sqrt{\dfrac{2|\Delta_m|}{\beta}} \tan(\eta t)}{1 + A_0 \sqrt{\dfrac{\beta}{2|\Delta_m|}} \tan(\eta t)}, \ \Delta_m > 0, \qquad \text{(III.F-11)}$$

$$B(t) = \frac{B_0 \sec^2(\eta t)}{\left[1 + A_0 \sqrt{\dfrac{\beta}{2|\Delta_m|}} \tan(\eta t)\right]^2}, \ \Delta_m > 0, \qquad \text{(III.F-12)}$$

$$A(t) = \frac{A_0 + \sqrt{\frac{2|\Delta_m|}{\beta}} \tanh(\eta t)}{1 + A_0 \sqrt{\frac{\beta}{2|\Delta_m|}} \tanh(\eta t)}, \ \Delta_m < 0. \tag{III.F-13}$$

$$B(t) = \frac{B_0 \, sech^2(\eta t)}{\left[1 + A_0 \sqrt{\frac{\beta}{2|\Delta_m|}} \tanh(\eta t)\right]^2}, \ \Delta_m < 0, \tag{III.F-14}$$

$$\eta = \sqrt{\frac{\beta |\Delta_m|}{2}}. \tag{III.F-15}$$

$$\Delta_m \equiv \alpha B_0 - \frac{\beta}{2} A_0^2. \tag{III.F-16}$$

IV. ASSUMPTIONS AND SOLUTIONS

IV.A. Introduction

Any model can be expressed in informational symbology. In the case of Lanchestrian (and Osepovian) attrition theory, the models are commonly expressed in three parts:

- a typical statement of what the model describes,
- a pair (normally) of coupled differential equations (which imply a solution), and
- a set of assumptions.

This chapter then is a general overview of the models that comprise basic Lanchestrian attrition theory.

Although it may seem somewhat premature, much of the discussion in this chapter centers on the elementary nature of the attrition rate constants/functions. In Chapter V, we will initially introduce the Ironman Analyses which lay the groundwork for the relationship between the attrition differential equations and the attrition rate constants/functions. This early discussion is, however, important in initiating the understanding of the interdependence between the two parts of Lanchestrian attrition mechanics: the theory of the attrition differential equations, and the conjugate theory of the attrition rate constants/functions.

IV.B. Lanchester's Linear Law.

The Linear Lanchester Law describes combat between two forces. The rate of the attrition is given by the differential equations

$$\frac{dA}{dt} = -\alpha AB, \qquad \text{(IV.B-1)}$$

$$\frac{dB}{dt} = -\beta AB. \qquad \text{(IV.B-2)}$$

The state solution for these differential equations, derived in Chapter III, is

$$\alpha(B - B_0) = \beta(A - A_0), \qquad \text{(IV.B-3)}$$

the explicit time solutions of the differential equations are derived in that chapter as well.

We note here that a multiplicative increase in attrition rate constant/function is equivalent to a multiplicative increase in force strength. If, for example, blue has an attrition rate twice that of red's, then blue's force strength need only be more than half red's to force victory (in the sense of a conclusion). If technology is used to this end, then its influence is direct and more efficient (of the two cases, linear and quadratic).

IV.B.1. Linear Law Assumptions.

The assumptions associated with this law are (following Dolansky[1] and Karr[2]):

1.) The two forces A (for amber or red) and B (for blue) are engaged in combat.

2.) The units of the two forces are within weapons range of all units of the other side.

3.) The attrition rates are known and constant.

4a.) Each unit is aware of the general location of enemy units but is unaware of the effect of fire.

5a.) Fire is uniformly distributed over the area occupied by enemy units.

6.a) The occupied area remains constant, units redisperse within the area

or

4b.) Each unit is aware of the specific location of enemy units and the effect of fire is known, but enemy units are hard to attrit, or are few in number (i.e. Hard to find.)

5b.) Fire from surviving units is uniformly distributed against enemy units.

We shall examine some of the implications of these assumptions.

Assumption (1) is perhaps the simplest and, at once, the most crucial. It seems intuitive that the model will only apply if combat is actually occurring. What must be noted is that combat is not a continuous process - it tends to be punctuated. Care must be taken to apply the model only when combat actually occurs.

This naturally leads to the concept (example here Agincourt & from Men at War?) that attrition must be time dependent. Further, it leads to the idea of time scales of combat. As we shall examine in a later chapter, the accommodation of attrition rates between theory and actuality (history) depends on the time scale that we consider. If we are interested in the losses per day, many of the actual combat processes become hidden. Historical data for combat losses seldom are available at time scales below one day. At this time scale, the dynamics of target acquisition become less important. Attrition rates are dominated by the ratio of enemy units killed (per day) to friendly rounds fired.

From a mathematical standpoint, the Lanchester differential equations freely admit introduction of a scaling fraction which is the time interval divided by the total time in combat. Either the time or the attrition rates may be scaled with this factor. If, for example, a unit is actively engaged in combat for an hour in a day, the only attrition caused by that unit (and possible suffered by it as well - attrition does not necessarily have to be symmetric in time,) occurs during that hour. This attrition translates directly into an attrition rate (which has units of inverse time - per minute or per hour), which is valid while the unit is engaged in combat. If the unit were continuously engaged in combat during the entire day, then the total attrition of that unit would be described by integration of the appropriate differential equation over the whole day's time. This, however, is not the case since (by premise) combat occurs only during one hour of the day. The total attrition of enemy units by that unit occurs only during that hour. To reconcile this limited attrition period with a total day of warfare for this unit, we may introduce a scaling fact ξ (which in this case has the value 24 - the number of hours in a day). This factor may be viewed as multiplying the time (which transforms combat time into elapsed time) or dividing the attrition rate (which transforms the in-combat attrition rate into an effective (daily) attrition rate). As we shall see in a later chapter, this problem is largely alleviated by the introduction

of attrition rates which are sensitive to the presence of enemy forces (such as range and/or time dependent attrition rates,) if not the actual state of combat.

Dolansky includes in Assumption (1) that the units engaged are identical but notes that this holds for only the simplest of Lanchester "type" models - he goes on to elaborate heterogeneous force Lanchester "type" models, which we discuss briefly in Chapter ( ).

This restriction gives occasion to treat an interesting case which illustrates the impact of military doctrine on attrition as well as the fundamental Lanchestrian question of what constitutes a unit. Some years ago, the doctrine of the Soviet Army, supposedly as a result of poor tank gun accuracy was that a tank platoon (3 tanks at that time, in that type of unit) would engage a single target collectively. The platoon leader would select a target. All three tanks would then take aim and fire together at that target. The unit of Soviet tank forces at that time was thus a platoon.

The tank forces of the NATO powers at that time, for the purpose of comparison, acquired and fired as individuals. Firing doctrine for NATO did not prescribe any type of deliberate mass firings (except perhaps accidentally or at responsive command discretion). Thus, the unit of NATO tank forces could be presumed to be an individual tank.

The consequences of these two doctrines in terms of attrition rates (and their calculation) will be discussed in a later chapter. Still, this difference points up some of the difficulty which arises in determining what actually comprises a unit in a Lanchestrian sense.

This difficulty is further demonstrated by Assumption (2), that each unit be within weapon range of all units on the other side. If we consider the case of combat in line with edged weapons (the Roman legions and their foes comes to mind as an example), then the lethal range of a weapon (sword, and/or non-thrown spear) is 1-2 meters. If the linear density of troops is ~ 1 per meter, then the Lanchestrian theory would seem to apply at about the level of one soldier fighting with one soldier. (This also make old Douglas Fairbanks movies seem to be correct in a Lanchestrian sense!) The unit would thus be the individual soldier. Description of combat ala Lanchester under these circumstances would seem then to be violated. A more reasonable assumption (which appears to yield the same result) would be that there are always targets within weapons range of all units. If we adopt this assumption, then as long as Assumption (1) holds, the result is the desired one. In terms of our Roman example, if the enemy line is maintained, then each Roman soldier in the Roman front line (engaged in combat) has 2-3 targets in range (in the enemy front line.)

The interpretation that presents itself, however, is that some areal structuring

of the attrition process is necessary. This is supported if we examine the frontage of troops in combat as a function of time (Dupuy[3] - We examine this in a later chapter) and compare this to weapon ranges. This interpretation is consistent with our revised Assumption (2).

The third assumption, that the attrition rates are known and constant is also open to discussion. Dolansky states that the attrition rates are difficult to evaluate (see the earlier discussion in this section of the difficulty of time scale adjustment). In principle, attrition rate constants/functions are calculable using Bonder's Equation (Chapter XII) although difficult to verify historically. Further, attrition rates for those factors of greatest interest, new weapons (the result of either new technology or human inventiveness), and new doctrine, inherently cannot be verified in terms of history. (We invoke a tacit, invisible subassumption here that warfare experiments - we do not include training and operational exercises with troops because of their controlled nature - cannot be conducted for whatever moral,ethical, and/or budgetary reasons.)

In spite of these difficulties, if we accept the applicability of Bonder's theory of attrition rates (that acceptance being an obvious, but defended premise of this volume), then the assumption that the attrition rates are known is satisfied; the assumption that the attrition rates are constant is much more difficult to accept or defend. In general, weapons' performance are range dependent. Further, there is considerable reason to believe that attrition rates should be time dependent as well. From a mathematical standpoint, constant attrition rates permit simple, straight forward closed form solutions of the Lanchester differential equations. Beyond this, however, assumption of constant attrition rates seems inconsistent with much of what we know of combat. It seems, therefore that this assumption is necessary not for the applicability of the Lanchester differential equations, but of the simple closed form solutions.

Assumptions (4a), (5a) and (6a) are generally supportive of what we think of as non line-of-sight weapon systems units - generally classical artillery (post American War Between The States) - whose operation is dependent on target acquisition information from other units and, because of a variety of position and time uncertainties, have only general knowledge of the position of enemy units. This uncertainty is mollified somewhat by the areal lethality of the weapon. These assumptions tend to be associated with the Lanchester linear law and lead to its identification with indirect fire weapons. The other pair of assumptions (4b) and (5b) are supportive of line-of-sight weapon units under conditions where targets are hard to attrit. This pair of assumptions is generally not associated with the linear law in much of the literature, although classically, of course, Lanchester associated the linear law with ancient combat, which was entirely line of sight attrition (except perhaps for some siege weapons(?), which are a special case).

The combination of assumptions (1) and (4b) seem to conflict with Lanchester's identification of the linear law with ancient combat. The law is assumed valid only when units are engaged in combat, and for ancient personal weapons, this effectively means that the units are in contact. It is then difficult to reconcile how the enemy units could be hard to find. The answer, of course, is that the units are not, while in contact, hard to find, but rather that because the forces are in contact, the rate of engagement is dependent on the product of densities of the two forces. This situation is directly comparable to a chemical reaction where the rate of the reaction is dependent on the concentration (densities) of the two (in this case?) reactant chemicals (forces). This analogy will be even more usefully applied in a later chapter on attrition processes where we develop the model of attrition as a scattering process. (It is interesting to note, using this analogy, that this type of chemical reaction description is valid when the reactants are completely mixed, as in a solution. A different form occurs when the reaction only occurs at (or in the region of) an interface.) This, has significant impact on attrition theory interpretation if we pursue the analogy. If this rate form is valid when the reactants are mixed, then the implication is that the forces must be mixed as well. This occurs only in a melee situation. Is then the norm of applying the Lanchester Linear Law to ancient combat melee combat only? Are ordered forms of ancient combat, such as those practiced in the phalanx and the legion, not described by such? This is indeed so as we shall see when we look at alternate forms of attrition "laws" such as those of Osipov and Helmbold. Among other things, we shall see there that the form of the attrition differential equation depends on the structure of the forces engaged, and that as that structure changes, so does the form of the differential equations.

We shall further see, in this chapter and in the chapter on the calculation of direct fire attrition rates that the form of the attrition rate can be either linear law-like or square law-like - a (not completely) general form is a combination which we shall consider in detail in another later chapter. We note, however, that Lanchestrian (Osipovian?) attrition theorists in the Soviet Union seem to sometimes perform both linear and square law calculations and use the two calculations as a bounded envelope about the "real" answers. Although we shall defer consideration of the linear law as descriptive of direct fire/line-of-sight/ point attrition to a later section, we will consider here the more normal association of the linear law as descriptive of indirect fire/ non line-of-sight area attrition. We may see the form of the attrition differential equation directly if in a somewhat simplified manner. Recalling Equation (IV.B-1), we may think of the entire red (A) force as occupying some area L. The number of red units per area is A/L. If each blue (B) fire kills all of the A force in a given area d about the impact point of munition, then the number of A force that is killed per blue fire is

$$\frac{A\,d}{L}\,. \qquad \text{(IV.B-4)}$$

If each blue unit fires m times in a given regular interval of time (e.g. rounds per minute), and we ignore such factors as overlapping lethal areas, shots outside of the occupied area L, and any question of target acquisition or weapon down time, then the number of A force units that are killed per blue unit per time (t_{Bk}) is A d m

$$\frac{Adm}{L} . \tag{IV.B-5}$$

Since the number of blue units is B, then the number of A force killed per time is just

$$\frac{d\ m}{L} A\ B, \tag{IV.B-6}$$

which is the attrition differential equation where

$$\alpha = \frac{dm}{L}, \tag{IV.B-7}$$

is the attrition rate. The minus sign, of course, arises because the total number of the A force is decreasing. The resulting attrition differential equation is simply,

$$\frac{dA}{dt} = -\alpha AB, \tag{IV.B-8}$$

which is a linear law attrition differential equation.

This also leads us to an understanding of the meaning of the attrition rates α and β. We may see that α is the number of A force units killed (by B) per B unit per time, per A unit. As we have seen in the simple development just above, it is not really the total number of A force units which is important, but rather their (areal) density. In fact, this leads to a sometimes stated assumption of indirect fire attrition theory - that forces are continuously redistributed over time (during combat) to keep a (changing) but constant areal density. Unfortunately, the presence of the red force strength as a factor in the attrition differential equation sometimes leads to confusion since the area of the forces dispersion is usually embedded in the attrition rate constant/function. This confusion can be somewhat alleviated (especially if closed form solutions are not being developed), by rewriting Equation (IV.B-1) as

$$\frac{dA}{dt} = -\alpha' \rho_A B, \tag{IV.B-9}$$

where: ρ_A = (areal) density of A

$$= \quad A/L\text{, and}$$
$$\alpha' \quad = \quad \text{revised attrition rate constant/function.}$$

(A similar parallel set of machinations can be performed for the other half of the combat attrition process that which occurs to B. For brevity we leave that as an exercise to the reader). Equation (IV.B-9) is a quadratic law differential equation. In keeping with the quadratic law (alternate) assumptions (specifically assumption (6b),) if the ratio of force strength to area occupied (that is, ρ_A,) remains constant, a quadratic law differential equation describes the attrition process.

If we do not neglect the target acquisition time, we must introduce a simple search and acquisition model. We have already defined the area density of Red units as ρA ($\equiv$ A/LA.) Let us postulate a search process where each unit of the Blue force searches an area l_B at any given time, with a probability p_{Aa} of finding a red unit in l_B (if a unit is present; we shall consider the effect of false detections in a later chapter;) and that the Blue unit searches areas of size l_B at a rate v_B (number of areas per time - the area searched per time is just $l_B\ v_B$.) The area per Red unit is just $\rho_A^{-1} = L_A/A$. The time required for a Blue unit to have searched an area which contains a Red unit is thus

$$\frac{L_A}{A l_B \nu_B}, or \qquad (\text{IV.B-10})$$
$$(\rho_A l_B \nu_B)^{-1}.$$

Since there is a probability p_{Aa} of the Blue unit detecting the Red unit, the probable number of areas that the Blue unit must search to find a Red unit is increased by a factor p_{Aa}^{-1}. The search time then becomes

$$t_{Bs} = \frac{L_A}{A l_B \nu_B p_{Aa}}, \qquad (\text{IV.B-11})$$
$$= (\rho_A l_B \nu_B p_{Aa})^{-1}.$$

We have earlier defined the time to kill (t_{Bk}) as

$$t_{Bk} = \rho_A dm. \qquad (\text{IV.B-12})$$

The total time to attrit a Red unit, including search and acquisition time (using these simple models,) is just

$$t_{Battrit} = t_{Bs} + t_{Bk}, \tag{IV.B-13}$$

and the attrition rate is

$$\begin{aligned} \alpha &= t_{Batttrit}^{-1}, \\ &= (t_{Bs} + t_{Bk})^{-1}. \end{aligned} \tag{IV.B-14}$$

If the search time is much greater than the kill time (i.e. $t_{Bs} >> t_{Bk}$,) then we may ignore t_{Bk} in the above attrition rate, and the attrition rate has the (approximate) form

$$\alpha \simeq \rho_A l_B v_B P_{Aa}. \tag{IV.B-15}$$

Now if ρ_A is constant, then the situation which we described earlier, namely that of the quadratic law assumption (6b) being valid, and the resulting differential equation has the form

$$\frac{dA}{dt} = -\alpha B, \tag{IV.B-16}$$

which is a quadratic law attrition differential equation. If, on the other hand, ρ_A is not constant, then square law assumption (6a) is valid, and the resulting differential equation has the form

$$\frac{dA}{dt} = -\alpha^* AB, \tag{IV.B-17}$$

which is a linear law attrition differential equation, and where:

$$\alpha^* = \frac{l_B v_B p_{Aa}}{L_A}. \tag{IV.B-18}$$

We shall further consider the interrelationship of search and kill times in Chapter VII which deals with combined law differential equations and assumptions.

1. Dolansky, Ladislav, "Present State of the Lanchester Theory of Combat, Operations Research 12 344-358, 1964.

2. Karr, Alan F., "Lanchester Attrition Processes and Theatre Level Combat Models", pp 89-126 in Shubik, Martin, ed., **Mathematics of Conflict**, North Holland,

New York, 1983.

3. Dupuy, Trevor N., **Numbers, Predictions and War**, Bobbs-Merrill, New York, 1979.

IV.B.2 Linear Law State Solution

Even though α is the attrition rate constant/function for the attrition of the A force, in the state solution, it is associated with the B force. This association occurs because α may also be interpreted with the effectiveness of the B force in attriting A. Note that we may interpret αB as the number of A units killed per time per A unit. The quantity Δ_1, defined by

$$\Delta_1 = \alpha B_0 - \beta A_0, \tag{IV.B-19}$$

is the difference in kills per time between the two forces at the beginning of the engagement. In terms of the Lanchestrian concept of combat to a conclusion, this difference is the predictor of victory. If $\Delta_1 > 0$ then the blue (B) force generates a larger number of kills per time than does its foe, the red (A) force. In this case, if the combat is carried to a conclusion, then the blue force will be the victor with

$$B_{final} = B_0 - \frac{\beta}{\alpha} A_0, \tag{IV.B-20}$$

units remaining.

(We shall discuss the historical perspective of combat to a conclusion in Chapter XIII.)

If $\Delta_1 < 0$, then the red (A) force generates more kills per time than does its foe, the blue (B) force. Thus, at the end of such a conclusive combat, the red force will be the victor with

$$A_{final} = A_0 - \frac{\alpha}{\beta} B_0, \tag{IV.B-21}$$

units remaining.

If $\Delta_1 = 0$, then the combat, if carried to a conclusion, results in a draw service both forces generate the same number of kills per time.

To examine the mathematical properties of the linear law, it is convenient to write the state solution, Equation (IV.B-3) in the form

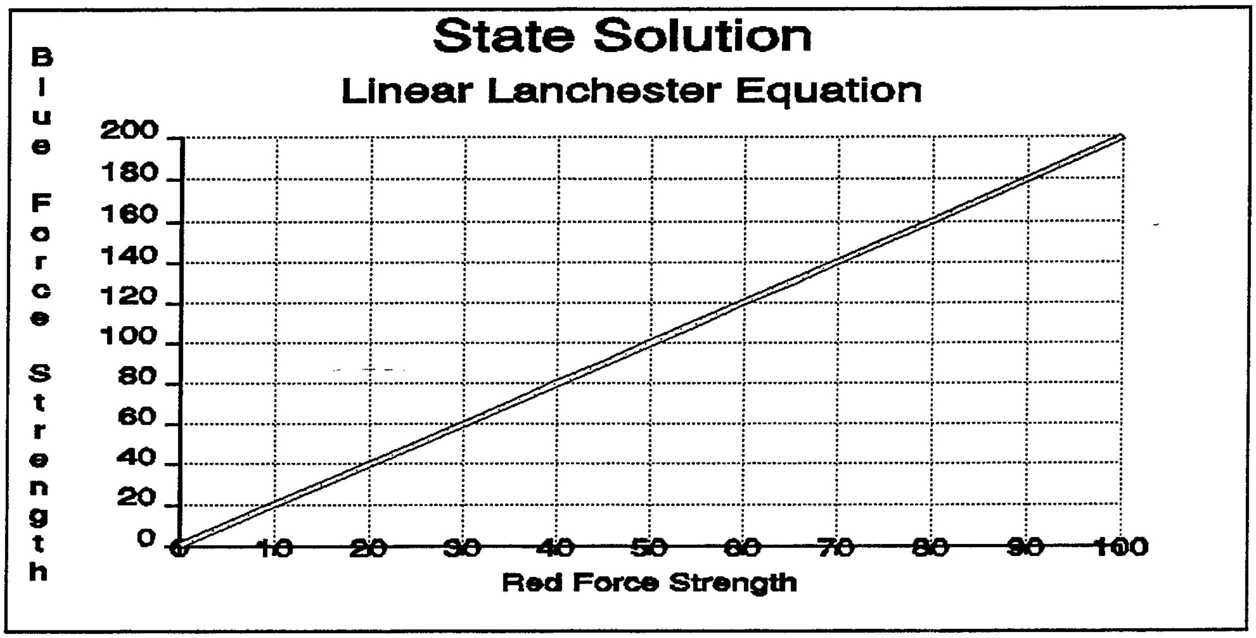
$$B = \frac{\beta}{\alpha} A + \frac{\Delta_1}{\alpha}. \tag{IV.B-22}$$

Mathematically, this is the equation of a straight line. (We have arbitrarily

chosen the red force strength (A) to serve as the independent variable while the blue force strength (B) serves as the dependent variable. Although we do this because

> (i) A comes before B in the alphabet, and
>
> (ii) we commonly associate blue with the friendly forces (except in the old Confederacy) and red with the enemy forces,

some convention needs to be established to provide a consistent basis for comparison. The reader is free to adopt the other convention, if desired, as an exercise). If we plot Equation (IV.B-23) in the normal manner with A = 0, B = 0 at



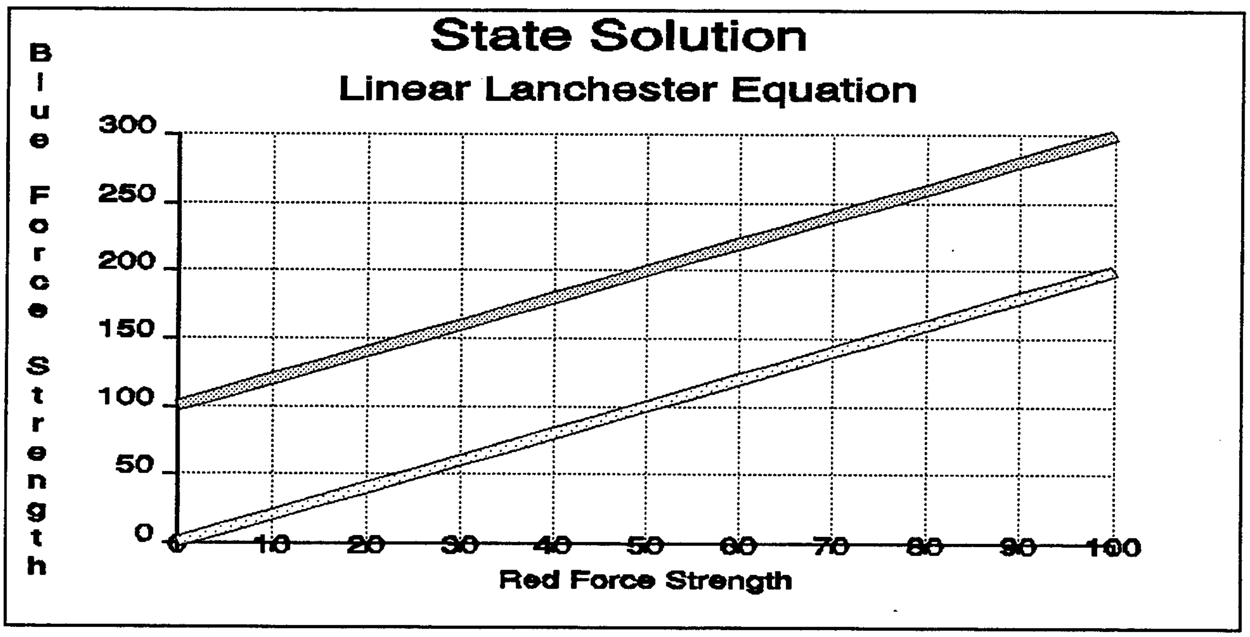
the origin of the axes (As shown in Figure (IV.B-1)), then we may see that the quantity β/α is the slope of the line.

$$B = \frac{\beta}{\alpha}A. \qquad \text{(IV.B-23)}$$

All solutions of the attrition differential equations (IV.A.1) and (IV.A.2) for these values of α and β (actually for this value of the ratio β/α ! in our convention) will lie parallel to this line. This line represents the case of a draw, when $\Delta_1 = 0$). This line also divides the graph into two regions, an upper and a lower region. The upper region contains those combats where the blue force is victorious (in the sense of conclusion), where $\Delta_1 > 0$. The lower region contains those combats where the red

force is victorious, where $\Delta_1 < 0$.

We may now examine combats in terms of the intercepts of the solutions with the axis. In the upper region, the state solution line must intercept the B force strength axis at zero A force strength. It does so at value Δ_1/α. We see now another interpretation of Δ_1; it is the number of kills per time remaining to the victor at the end of a conclusive combat; it "represents" the power or ability of the victorious force to enter further combat. Further, divided by the appropriate attrition rate constant/ function, Δ_1 is the force strength of the victor at the conclusion of combat, (Note that in our convention, a plus sign here indicates a Blue force victory; a minus sign indicates a Red force victory). These cases are shown in Figures (IV.B-2) and (IV.B-3) respectively. The values of α and β are held constant (and equal). In Figure (IV.B-2), the initial Blue force strength is increased by 50%. Note that this 50% is the entirety of the Blue force remaining at the conclusion. (The graph is read in a right to left manner. The battle begins at the upper right hand edge [above the draw line], and proceeds down and to the left). In figure (IV.B-3), the initial Blue force strength is 50% less than in the draw case. Note that 50% of the Red force remains at conclusion. This points up one way to win a victory (under conclusion conditions,) the side with the larger force (numbers) wins.



Another way to win is to change the attrition rate constants/ functions. Recall the (simple) definition (model) of the attrition rate constants/functions.

$$\alpha = \frac{d_B m_B}{L_A}, \tag{IV.B-24}$$

and

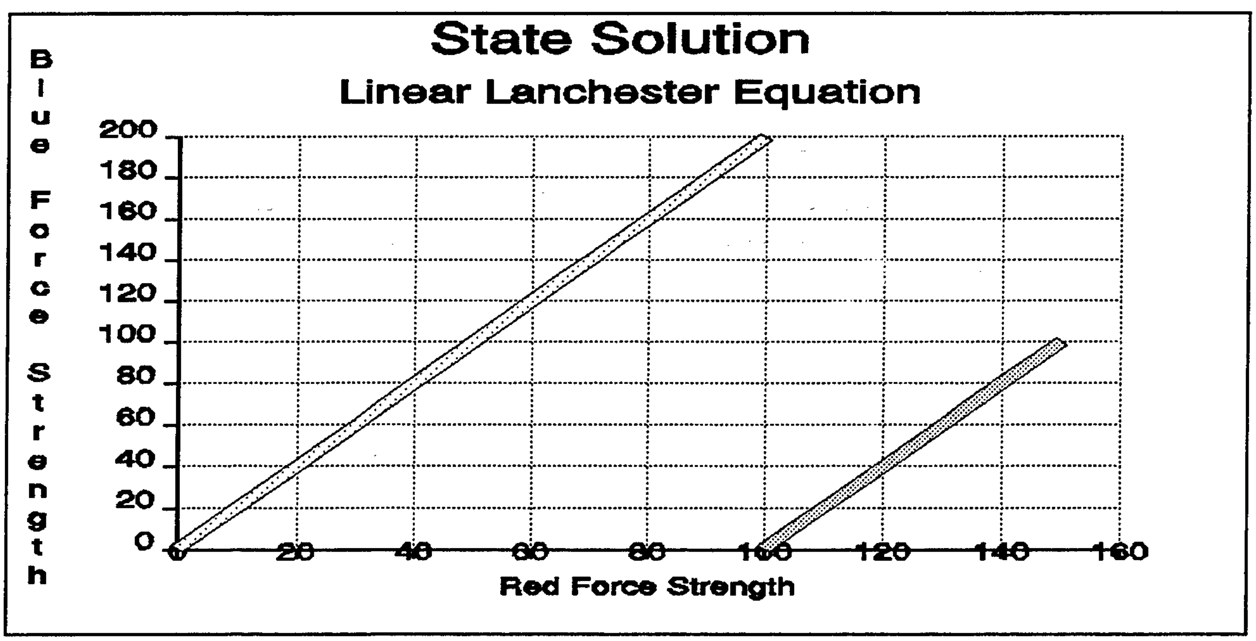
$$\beta = \frac{d_A m_A}{L_B}, \tag{IV.B-25}$$

where: d_A, d_B = lethal area of A, B force shot,
m_A, m_B = rate of fire of A, B unit, and
L_A, L_B = area occupied by A, B forces.

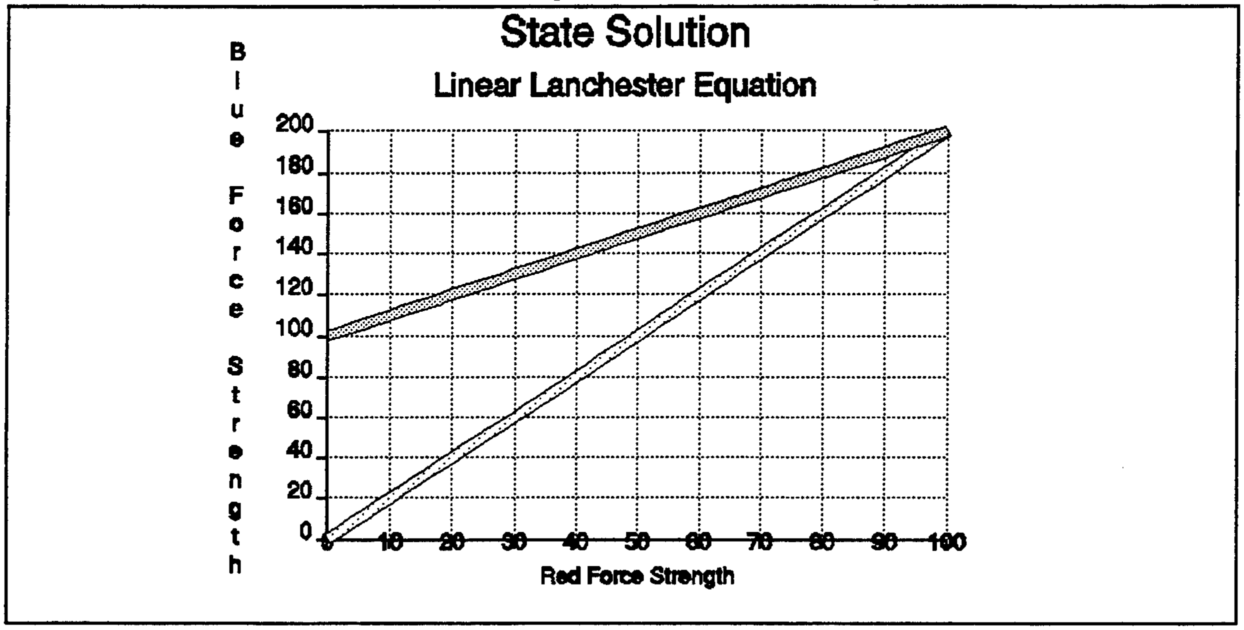
There are basically three ways to change the attrition rate model. We shall examine each of these in turn holding the initial force strengths of both forces fixed at the values in the draw case, and holding fixed the three parameters:

- lethal area per fire,
- rate of fire per unit, and
- occupied area of the Red force.

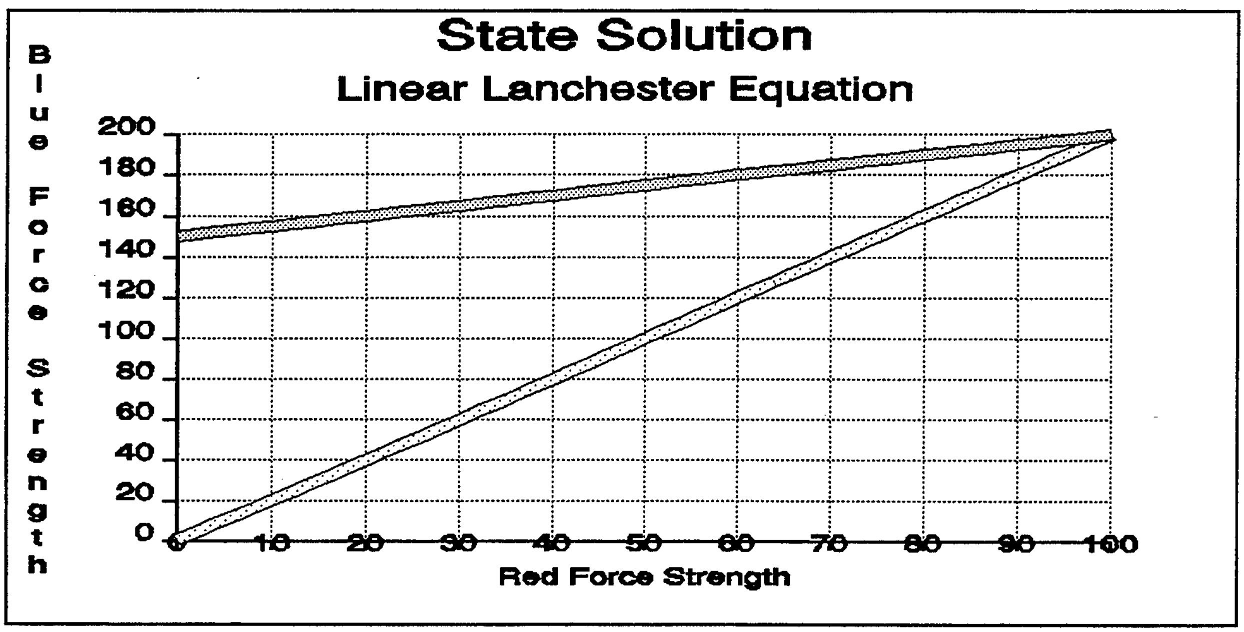
The first way to change the attrition rate constants/functions is to change the area occupied by the Blue force. This has no effect on the rate of attrition of the Red force. Rather, it decreases the number of Blue units struck by each Red unit fire - it decreases the rate of attrition of the Blue force. In other words, if we double the area



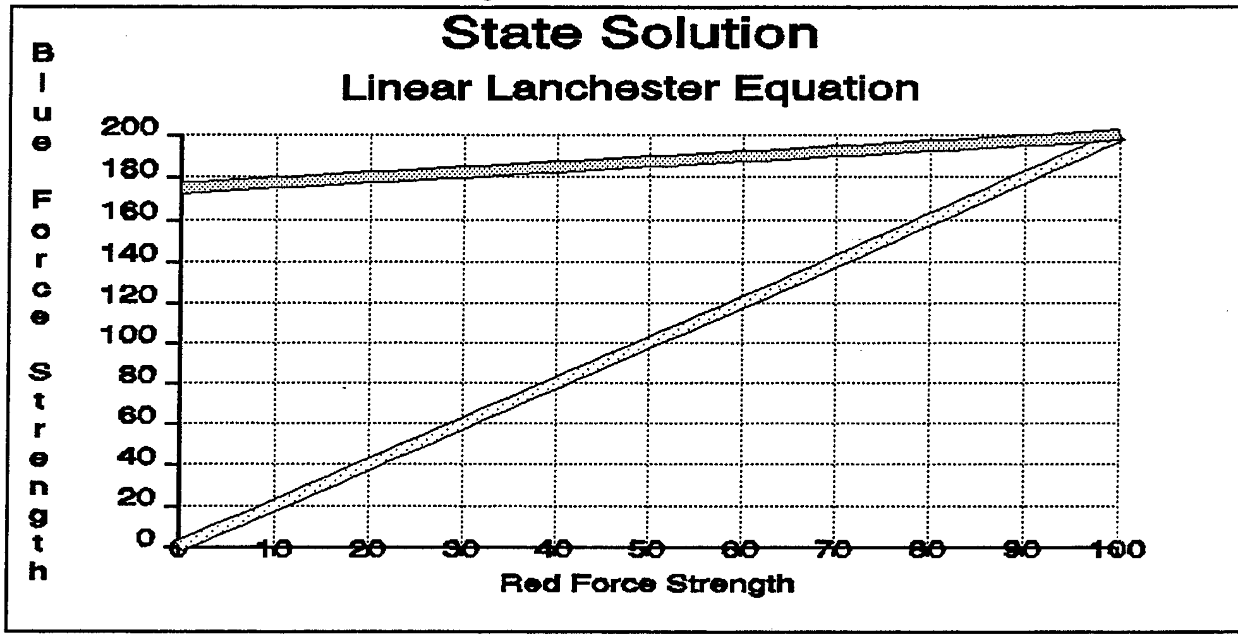
that the Blue force occupies, α becomes ½ of its previous value. This case is plotted in Figure (IV.B-4). This change is most likely doctrinal in nature, assuming the Blue force's infrastructure, such as Command, Control, and Communication, can support the dispersal. Notice that dispersing the force this way may violate Assumption (2) since all of the Blue force may no longer be in weapons range of all of the Red force.



The second and third ways to change the attrition rate constants/functions are for Blue (in this case) to increase the lethal area of his munitions and/or to increase



his rate of fire. Lethal area can be increased achieved by adopting larger weapons/munitions (which usually decreases the rate of fire), or by technological improvement of the munition (such as better explosives). The rate of fire can be increased by training the weapon crews better, or again by technological improvements, such as by incorporating automatic loading. If we double either of these parameters, β doubles over its draw case value while α stays the same. This case is shown in Figure (IV.B-5). (Note: This is identical to Figure (IV.B-4).) If we double both parameters, β quadruples over its draw case value while α stays the same. This case is shown in Figure (IV.B-6).



These investigations display the general characteristics of indirect fire combat as described by Lanchester's linear law:

- Maximum force dispersion, consistent with weapon effectiveness minimizes losses. (We note in passing that this is also the case when direct fire attrition is described by the linear law. It also applies to the use of weapons of mass destruction - nuclear and chemical weapons).

- Increased weapons effectiveness decreases casualties.

- Economy of force is manifested in the use of minimum force strength to effect the mission (casualties are linear).

IV.B.3. Linear Law Time Solution

The time solution of the Lanchester linear attrition differential equation , derived in Chapter III are

$$A(t) = A_0 \frac{\Delta_1}{\beta A_0 - \alpha B_0 e^{-\Delta_1 \Delta t}}, \tag{IV.B-26}$$

and

$$B(t) = B_0 \frac{\Delta_1 e^{-\Delta_1 \Delta t}}{\beta A_0 - \alpha B_0 e^{-\Delta_1 \Delta t}}, \tag{IV.B-27}$$

where:

$$\Delta_1 \equiv \beta A_0 - \alpha B_0. \tag{IV.B-28}$$

We note immediately that we cannot obviously solve these equations for a draw case - both equations (IV.B-26) and (IV.B-27) appear to be zero when $\Delta_1 = 0$. (The general case of draw solutions are considered in Chapter VI.) They can however, be solved for $\Delta_1 \neq 0$. The draw case can be considered if we expand the exponential terms in Equations (IV.B-26) and (IV.B-27) to first order in Δ_1,

$$e^{\pm \Delta_1 \Delta t} \simeq 1 \pm \Delta_1 \Delta t, \tag{IV.B-29}$$

which we substitute into those two equations (after we rearrange equation (IV.B-27) to have only one exponential term. This yields

$$A(t) = A_0 \frac{\Delta_1}{\beta A_0 - \alpha B_0 (1 - \Delta_1 \Delta t)}, \tag{IV.B-30}$$

and

$$B(t) = B_0 \frac{\Delta_1}{\beta A_0 (1 + \Delta_1 \Delta t) - \alpha B_0}, \tag{IV.B-31}$$

which reduces, using the definition of Δ_1 to

$$A(t) = \frac{A_0}{1 + \alpha B_0 \Delta t}, \tag{IV.B-32}$$

and

$$B(t) = \frac{B_0}{1 + \beta A_0 \Delta t}. \tag{IV.B-33}$$

To calculate particular solutions of these equations, we must first compute values of α and β, and assume some initial force strengths. As examples we take,

$A_0 = 100$, and
$B_0 = 200$.

The attrition rates, in the simplest case of kill dominated attrition, are

$$\alpha = \frac{d_B m_B}{L_A}, \tag{IV.B-34}$$

and

$$\beta = \frac{d_A m_A}{L_B}, \tag{IV.B-35}$$

where: d_A, d_B = lethal area of A, B force shot,
m_A, m_B = rate of fire of A, B unit, and
L_A, L_B = area occupied by A, B forces.

For Δ_1 to be zero, β must be twice α. We take initially then,

LA = 100 km^2,
$d_B = 1\ km^2$, and
$m_B = 5\ min^{-1}$.

This gives $\alpha = 5 \times 10^{-2}\ min^{-1}$. If we take $L_B = L_A$, $m_A = m_B$, and $d_A = 2\ km^2$, then $\beta = 10^{-1}\ min^{-1}$. This satisfies the draw case condition. A plot of equations (IV.B-32) and (IV.B-33) for these parameters are given in Figure (IV.B-7).

Variations for doubled/halved force strengths and doubled occupation area, rate of fire/lethal area are shown in subsequent figures. Since these examples deviate from the draw case, the force strengths were calculated using Equations (IV.B-26) and (IV.B-27). Note how the draw shifts to Blue/Red victory in a conclusion sense.

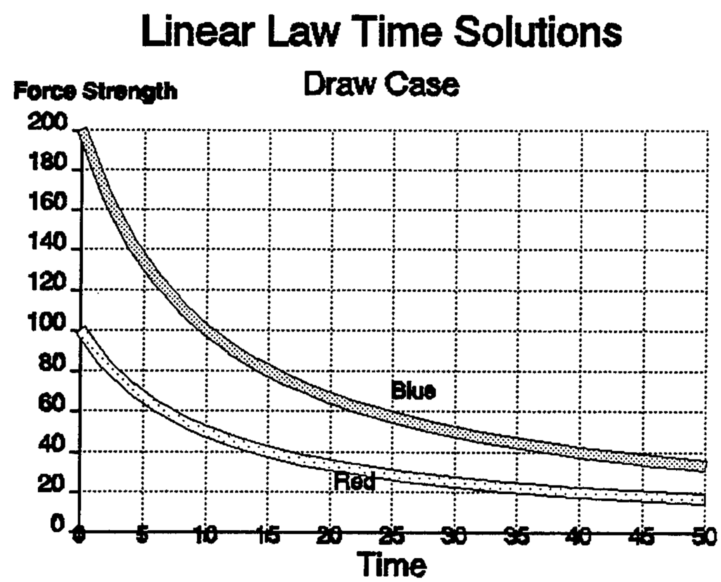


Figure IV.B-7

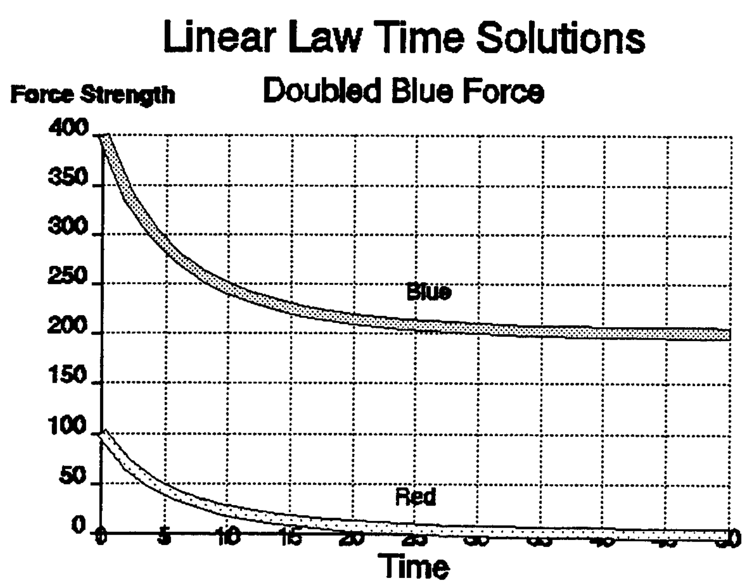


Figure IV.B-8

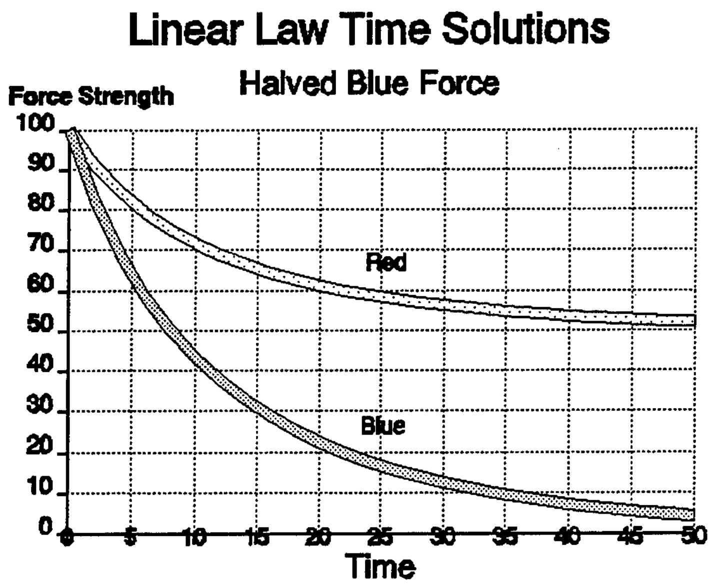


Figure IV.B-9

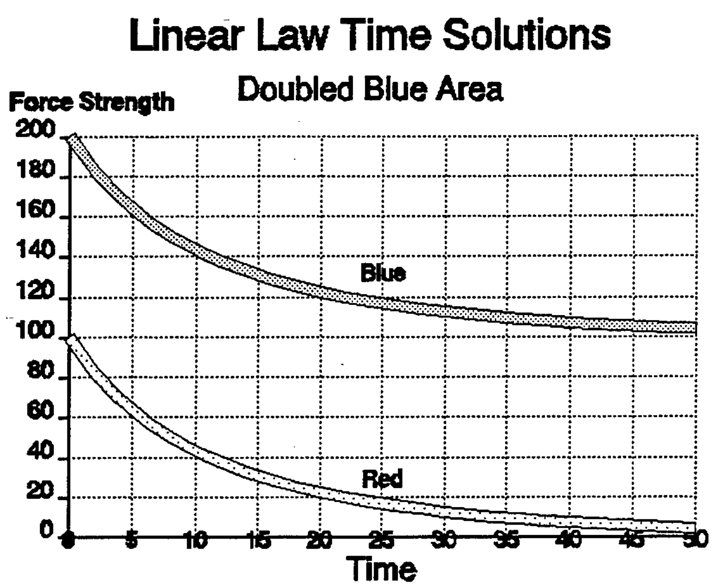


Figure IV.B-10

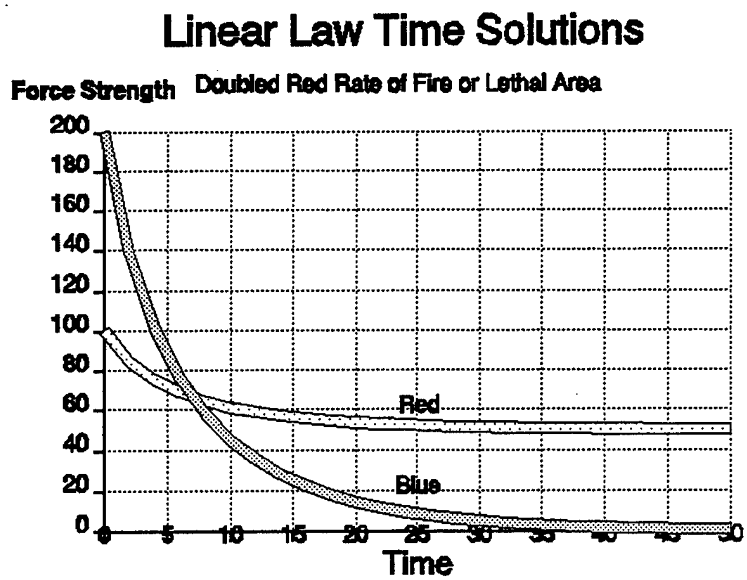


Figure IV.B-11

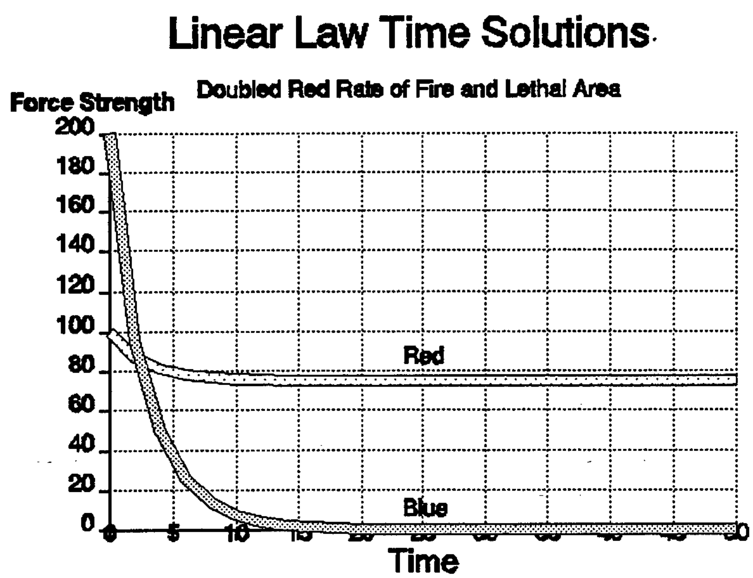


Figure IV.B-12

IV.C. Lanchester's Square Law

The square Lanchester law also describes combat between two forces. The rate of attrition is given by the differential equations

$$\frac{dA}{dt} = -\alpha B, \tag{IV.C-1}$$

and

$$\frac{dB}{dt} = -\beta A. \tag{IV.C-2}$$

The state solution for these differential equation, derived in Chapter III, is

$$\alpha(B^2 - B_0^2) = \beta(A^2 - A_0^2). \tag{IV.C-3}$$

The explicit time solutions of these differential equations are derived in that chapter as well.

In the square law case, as in the linear law case, an increase in attrition rate constant/function is equivalent to a multiplication increase in force power. (Such an increase in attrition rate constant/function increases force strength only as the square root since force power (total force kills per time) is the attrition rate constant/function times the square of the force strength rather than as the force strength directly in the linear law). If for example, Blue has an attrition rate constant/function twice Red's, then Blue's force strength need only be slightly more than 70% of Red's force strength to force victory (again, in the sense of a conclusion). If technology is used to this end, then its influence is still direct, but is less efficient (of the two cases) since the attrition rate constant/function must quadruple for every factor of two that the enemy force strength increases.

This is a direct statement of Lanchester's Principle of Concentration.

IV.C.1. Square Law Assumptions

The assumptions associated with the square law are, again following Dolansky, and Karr:

1.) The two forces A (for amber or red) and B (for blue) are engaged in combat.

2.) The units of the two forces are within weapons range of all units of the other side.

3.) The attrition rates are known and constant.

4a.) Each friendly unit is aware of the specific location of enemy units and the effect of fire is known.

5a.) Fire is uniformly distributed over surviving enemy units.

6.a) Targets are either numerous or are acquired at a constant rate (i.e. are easy to find.)

or

4b.) Each friendly unit is aware of the general location of enemy units but the effect of fire is generally unknown.

5b.) Fire from surviving friendly units is uniformly distributed over the area occupied by enemy units.

6b.) The area occupied by surviving units contracts to maintain a constant density of units.

We notice immediately that the first three assumptions, (1)-(3), are the same as those advanced for the linear law. The reader is referred to the previous section for discussion of those assumptions. We shall concentrate here on the "new" assumptions which apply to the square law.

Assumptions (4a), (5a) and (6a) are those commonly associated with the square law as a model of line-of-sight weapon systems units - generally classical infantry and cavalry/armor units, and artillery units firing directly. (Artillery units were predominantly direct fire until after the period of the American Civil War/War of Southern Independence circa 1861-1865 C.E.) These assumptions describe direct fire combat when targets are easy to find and the attrition rate process is dominated directly by the rate of fire/kill rather than by the target location/identification process. (As described by assumptions (4b) and (5b) of the linear law). Assumptions (4a) - (6a) are those which we have seen support indirect fire combat and the comments in the previous section are still applicable, but are modified by assumption (6b). In this case, the quantity

$$\rho_A = \frac{A_0}{L_A}, \qquad \text{(IV.C-4)}$$

(and its conjugate) are conserved through the combat. As a result, the area occupied by each force L_A, L_B , are now time dependent, and have the form

$$L_A = \rho_A A, \tag{IV.C-5}$$

so that the attrition rate has the form

$$\begin{aligned} \alpha &= \frac{d_B m_B}{L_A} AB \\ &= \frac{d_B m_B}{\rho_A A} AB \\ &= \frac{d_B m_B}{\rho_A} B, \end{aligned} \tag{IV.C-6}$$

and the square form of the Lanchester differential equations arise. The indirect fire attrition rate constant/function for constant density of forces is related to that for constant area occupied by forces (designated by and α_ρ and α_A, respectively) is

$$\alpha_\rho = \alpha_A L_A, \tag{IV.C-7}$$

where L_A here is the area occupied by the initial forces.

If we again consider the search and acquisition time in the attrition rates, the search model previously described in section IV.B may be used. The search time is again

$$\begin{aligned} t_{Bs} &= \frac{L_A}{A l_B v_B p_{Aa}} \\ &= (\rho_A l_B v_A p_{Aa})^{-1}. \end{aligned} \tag{IV.C-8}$$

The kill time is just

$$t_{Bk} = (rp)^{-1}, \tag{IV.C-9}$$

where r is the rate of fire of the weapon, and p is the probability of kill per shot. The total time to attrit a Red unit, including search and acquisition time (using these simple models,) is just

$$t_{Battrit} = t_{Bs} + t_{Bk}, \tag{IV.C-10}$$

and the attrition rate is (again)

If the search time is much greater than the kill time (i.e. $t_{Bs} >> t_{Bk}$,) then we may ignore t_{Bk} in the above attrition rate, and the attrition rate is again,

$$\alpha = t_{Battrit}^{-1} = (t_{Bs} + t_{Bk})^{-1}. \tag{IV.C-11}$$

$$\alpha = \rho_A \, l_B \, v_B \, p_{Aa}. \tag{IV.C-12}$$

Now if ρ_A is constant, then quadratic law assumption (6b) is valid, and the resulting differential equation has the form

$$\frac{dA}{dt} = -\alpha \, B, \tag{IV.C-13}$$

which is a quadratic law attrition differential equation (regardless of the type of attrition.) If, on the other hand, ρ_A is not constant, then square law assumption (6a) is valid, and the resulting differential equation is a linear law attrition differential equation,

$$\frac{dA}{dt} = -\alpha^* \, A \, B, \tag{IV.C-14}$$

where:

$$\alpha^* = \frac{l_B \, v_B \, P_{Aa}}{L_A}. \tag{IV.C-15}$$

If the kill time is much greater than the search time (i.e. $t_{Bs} << t_{Bk}$,) then we may ignore t_{Bs} in the above attrition rate, and the attrition rate has the (approximate) form

$$\alpha = r \, p, \tag{IV.C-16}$$

and the resulting attrition differential equation is quadratic.

The interrelationship of search and kill times will be further considered in Chapter VII which deals with combined law differential equations and assumptions.

On a historical basis, one would expect the actuality of combat to 'see-saw' between the linear and square law descriptions of indirect fire combat. Initially, units would be distributed over an area and would remain so for some time. Then, casualties having occurred in a non-uniform manner, the surviving units might be redistributed (over a lesser area) to fill gaps but reverting to approximately their

original density. During the period of redistribution, we might expect that the square law form would hold. This view, of course, is somewhat simplistic (but no more so than the model itself). It will also depend on whether combat is continued (and to what intensity,) while the units are redistributed. Alternatively, the area occupied by the forces will tend to remain somewhat constant even when casualties occur due to the need to maintain a force presence in those areas. This is a subject that we shall also take up in Chapter VII.

It is worth commenting that one of the assumptions in the Lanchester model describing indirect fire units (or those affected by indi-rect fire) is that such units are uniformly distributed. This is only approximately so. The individual weapon systems may be approximately uniformly distributed over an area (or a line) with some degree of concentration, but by their very nature, the portions of the force which are not (usually) attriters, (i.e., command and supply units,) by their very nature are concentrated and not so distributed. The model is too simplistic (at this level of development and discussion) to consider these units or the effects of their attrition. This concentration is why target location has become crucial for indirect fire systems - the need to selectively engage these control and support units which are not efficiently attrited under the normal assumptive conditions.

IV.C.2. Square Law State Solution

To consider the square law as descriptive of direct fire / line-of-sight/point attrition, we again perform a simple analysis. Consider that each Blue fire is directed against one Red unit at a time (assuming a unit to be the simplest level of weapon system, such as a tank or an individual soldier. If the unit is larger - a squad or platoon, say - then this condition still applies but the unit attrition is fractional. We shall illustrate this later when we analyze the example of the Soviet tank platoon as unit). Associated with each unit is a rate of fire (fires per time) of r_u and a probability of kill per shot of p_u. If target location/identification time is small compared to time to kill once the target is located (a situation dictated by assumptions (4b) and (5b)), and the target unit is engaged until killed (and (!) ammunition supply is ignored), then the time to kill a Red unit is just $(r_u\, p_u)^{-1}$, and the attrition rate is just

$$\alpha = r_u\, p_u, \tag{IV.V-17}$$

which is yields a linear attrition differential equation.

If we again define the quantity Δ_2 as

$$\Delta_2 \equiv \alpha B_0^2 - \beta A_0^2, \tag{IV.V-18}$$

which is the kills per time difference between the two forces. As with Δ_2, this is the predictor of victory in the Lanchestrian sense of combat to a conclusion. As before, if $\Delta_2 > 0$, then the blue force generates more kills per time than does its foe, and if combat is carried to a conclusion, then the blue force will be the victor with

$$B_{final} = \sqrt{B_0^2 - \frac{\beta}{\alpha} A_0^2} \tag{IV.V-19}$$

units remaining.

If $\Delta_2 < 0$, then the red force generates more kills per time than does the blue force, and at the end of a conclusive combat, the red force will be the victor with

$$A_{final} = \sqrt{A_0^2 - \frac{\alpha}{\beta} B_0^2} \tag{IV.V-20}$$

units remaining.

If $\Delta_2 = 0$, then the combat, if carried to a conclusion, results in a draw since both forces generate the same number of kills per time.

It is again convenient to write the state solution in a form when the red force strength is the dependent variable and the blue force strength is the independent variable,

$$B = \sqrt{\frac{\beta}{\alpha}A^2 + \frac{\Delta_2}{\alpha}}. \qquad \text{(IV.V-21)}$$

If we plot this function for $\Delta_2 = 0$, we get a graph of the same form as Figure (IV.B-1), except that the slope is

$$\sqrt{\frac{\beta}{\alpha}}. \qquad \text{(IV.V-22)}$$

This follows since

$$B = \sqrt{\frac{\beta}{\alpha}}\,A, \qquad \text{(IV.V-23)}$$

when $\Delta_2 = 0$

Since equation (IV.C-21) is quadratic rather than linear, solutions for various combats will not lie parallel to this line (for the same ratio α/β), as they did in the linear case. They will, however, lie either above or below this line, respectively, whenever $\Delta_2 < 0$, or $\Delta_2 > 0$. As before, if $\Delta_2 > 0$, the solution will graph above this line and Blue will be victorious (in a conclusive combat). If $\Delta_2 < 0$, the solution will graph below the line and Red will be victorious.

We now examine, in the same manner as previously, combats in terms of the intercepts of the solutions with the axis. In the upper region, the state solution curve intercepts the B force strength axis at zero A force strength. It does so at value $\sqrt{(\Delta_2/\alpha)}$. The quantity Δ_2/α again represents the number of kills per time remaining to the victor at the end of a conclusive combat. The quantity $\sqrt{(\Delta_2/\alpha)}$ is the force strength of the victor at the conclusion of combat.

As before, we examine the effect of force strength on the outcome of the battle. This is shown in Figures (IV.C-1) and (IV.C-2) for an increase and a decrease in the initial Blue force strength of 50% over the draw case. Note that in the latter case, a Red victory, the Red force strength at conclusion, the interaction of the state solution curve with the A force strength axis, is $\sqrt{(-\Delta_2/\beta)}$. In the square law case, we must worry about the sign of Δ_2 explicitly since the argument of the square root must be positive. These curves show that one way for Blue to win is to have more units

than Red (for the same attrition rate constant/function).

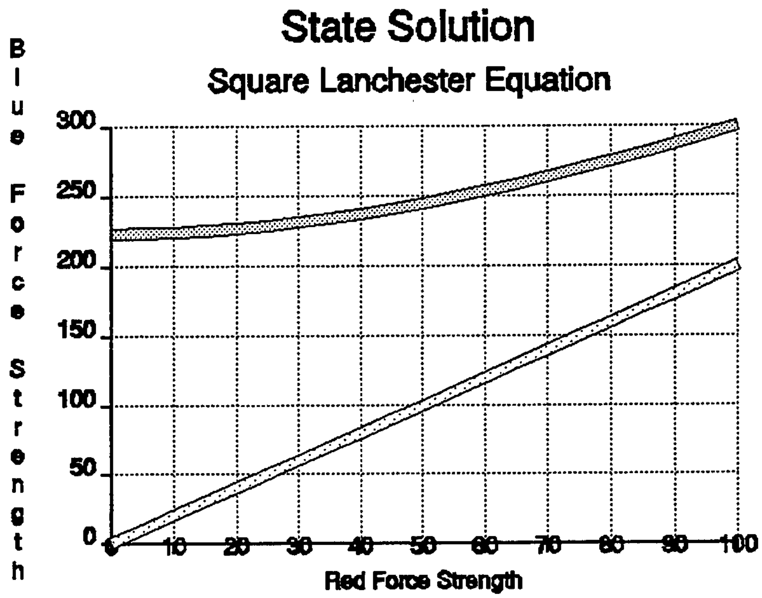


Figure IV.C-1

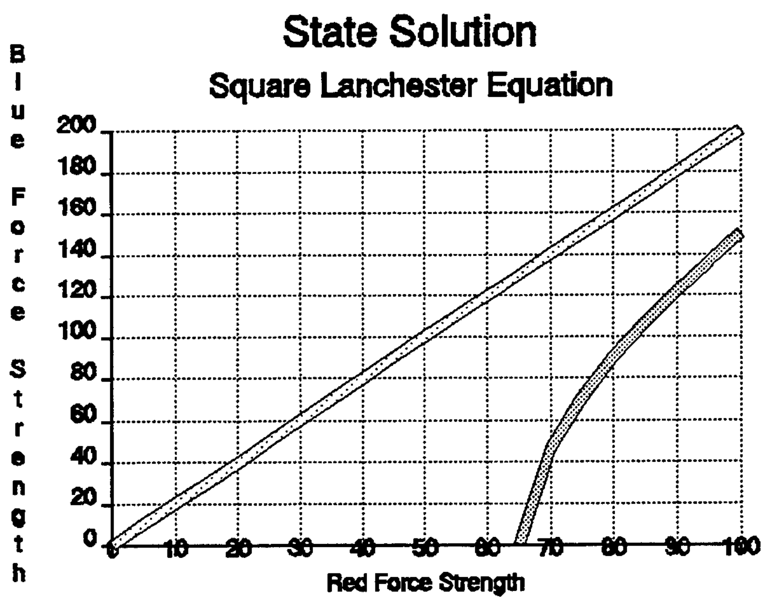


Figure IV.C-2

The other way for Blue to win is to change the attrition rate constant/function. This cannot be done by increasing the area occupied as in the linear case; it can only be done by increasing either the rate of fire or the lethality of the munitions (increasing

the probability of kill). If we double either the probability of kill or the rate of fire, the Blue attrition rate is doubled in value (com-pared to the draw case). The result of this is shown in Figure (IV.C-3). If we double both, the Blue attrition rate is quadrupled in value. This result is shown in Figure (IV.C-4). As in the square law case, the attrition rate can be changed through either training or technology. Both rate of fire and probability of kill can be increased by developing the skills of the loader (assuming a manual loader,) or the skills of the gunner, respectively. Similarly, by incorporating an automatic loader (increasing the rate of fire when the unit is kill limited,) or improving the accuracy of the weapon and/or the lethality of the munition (increasing probability of kill,) the attrition rate can be increased.

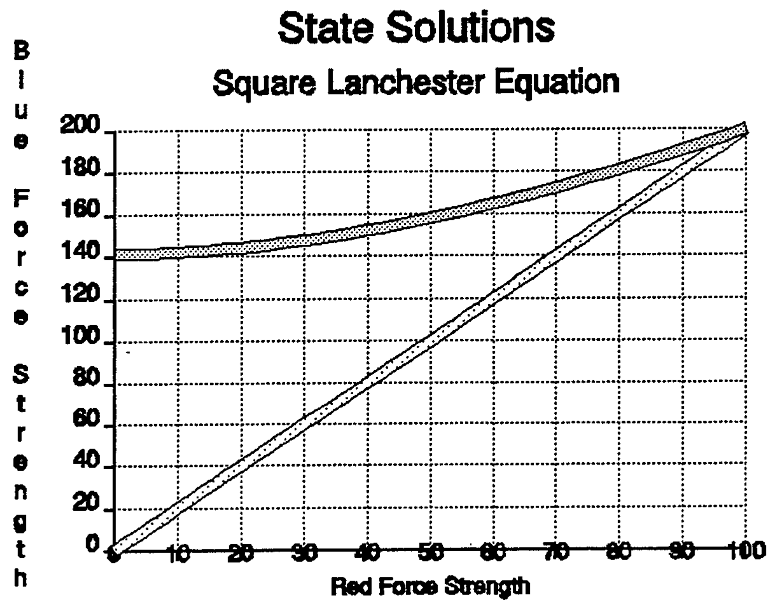


Figure IV.C-3

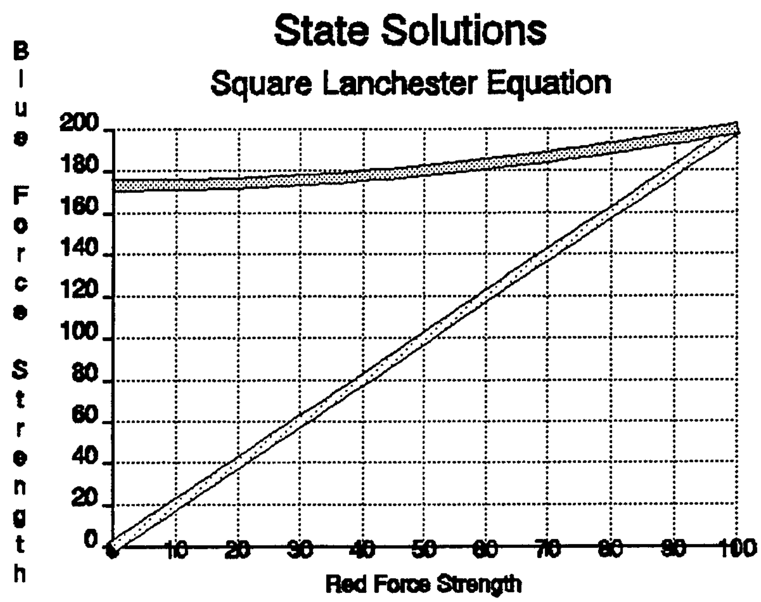

State Solutions
Square Lanchester Equation
Blue Force Strength
200
180
160
140
120
100
80
60
40
20
0
10
20
30
40
50
60
70
80
90
100
Red Force Strength

Figure IV.C-4

IV.C.3. Square Law Time Solution

The square law time solutions, derived in Chapter III, are:

$$A(t) = A_0 \cosh(\gamma t) - \delta\, B_0 \sinh(\gamma t),$$
$$B(t) = B_0 \cosh(\gamma t) - \frac{A_0}{\delta} \sinh(\gamma t), \tag{IV.C-24}$$

where:

$$\Delta_2 = \alpha B_0^2 - \beta A_0^2,$$
$$\gamma = \sqrt{\alpha\beta}, \tag{IV.C-25}$$
$$\delta = \sqrt{\frac{\alpha}{\beta}}.$$

It is not obvious that these equations are valid for the draw case. To show this, we first rewrite Δ_2 in the form,

$$\Delta_2 = (\sqrt{\alpha} B_0 - \sqrt{\beta} A_0)\,(\sqrt{\alpha} B_0 + \sqrt{\beta} A_0). \tag{IV.C-26}$$

We see that the draw condition $\Delta_2 = 0$ means that

$$\sqrt{\alpha} B_0 - \sqrt{\beta} A_0 = 0, \tag{IV.C-27}$$

or

$$\sqrt{\alpha} B_0 = \sqrt{\beta} A_0. \tag{IV.C-28}$$

(This is also the result that we would have gotten if we had solved Equation (IV.C-25) directly.)

If we now consider the alternate solution forms in Appendix C:

$$2\sqrt{\beta} A(t) = (\sqrt{\beta} A_0 - \sqrt{\alpha} B_0) e^{\gamma \Delta t} + (\sqrt{\beta} A_0 + \sqrt{\alpha} B_0) e^{-\gamma \Delta t}, \tag{IV.C-29}$$

and

$$2\sqrt{\alpha} B(t) = (\sqrt{\alpha} B_0 - \sqrt{\beta} A_0) e^{\gamma \Delta t} + (\sqrt{\alpha} B_0 + \sqrt{\beta} A_0) e^{-\gamma \Delta t}, \tag{IV.C-30}$$

and substitute Equation (IV.C-27) into these equations, we obtain,

$$2\sqrt{\beta}A(t) = (\sqrt{\beta}A_0 + \sqrt{\alpha}B_0)e^{-\gamma\Delta t}, \tag{IV.C-31}$$

and

$$2\sqrt{\alpha}B(t) = (\sqrt{\alpha}B_0 + \sqrt{\beta}A_0)e^{-\gamma\Delta t}, \tag{IV.C-32}$$

And now substitute Equations (IV.C-28) into these two equations, and perform some minor algebra,

$$\begin{aligned} A(t) &= A_0\, e^{-\gamma\Delta t} \\ B(t) &= B_0\, e^{-\gamma\Delta t}, \end{aligned} \tag{IV.C-33}$$

result. (We shall derive these equations from the attrition differential equations in Chapter VI.)

For an example of the draw case, we again take

$A_0 = 100$, and

$B_0 = 200$.

The ratio β/α must be 4. From the simple model of attrition,

$$\begin{aligned} \alpha &= r_B\, p_B \\ \beta &= r_A\, p_A, \end{aligned} \tag{IV.C-34}$$

we see that this may be satisfied if:

(i.)	$r_A = 4\ r_B$
(ii.)	$p_A = 4\ p_B$
(iii.)	$r_A = 2\ r_B$, and $p_A = 2\ p_B$
(iv.)	$r_A = 2\ x\ r_B$, and $p_A = 2\ p_B/x$

where $x > 0$. For the purpose of this example, we will take case (iii.) above, and use

$r_B = 3\ \text{min}^{-1}$, and

$p_B = 0.25$.

This results in $\alpha = 0.75\ \text{min}^{-1}$, $\beta = 3.00\ \text{min}^{-1}$, $\Delta_2 = 0.00$, $\gamma = 1.50\ \text{min}^{-1}$, and $\delta = 0.50$. A plot of these particular solutions for the draw case, Equations (IV.C-33) are given in Figure (IV.C-5). While the ultimate convergence of the two solutions ate force strengths of zero is not shown in this figure (in the interest of keeping a reasonable span on the chart,) that end is clearly indicated. As with the linear law conclusion condition, the square law conclusion condition can be changed in two ways, by changing the initial force strengths and by changing the attrition rate constants/ functions. Each of these variations is depicted in subsequent figures.

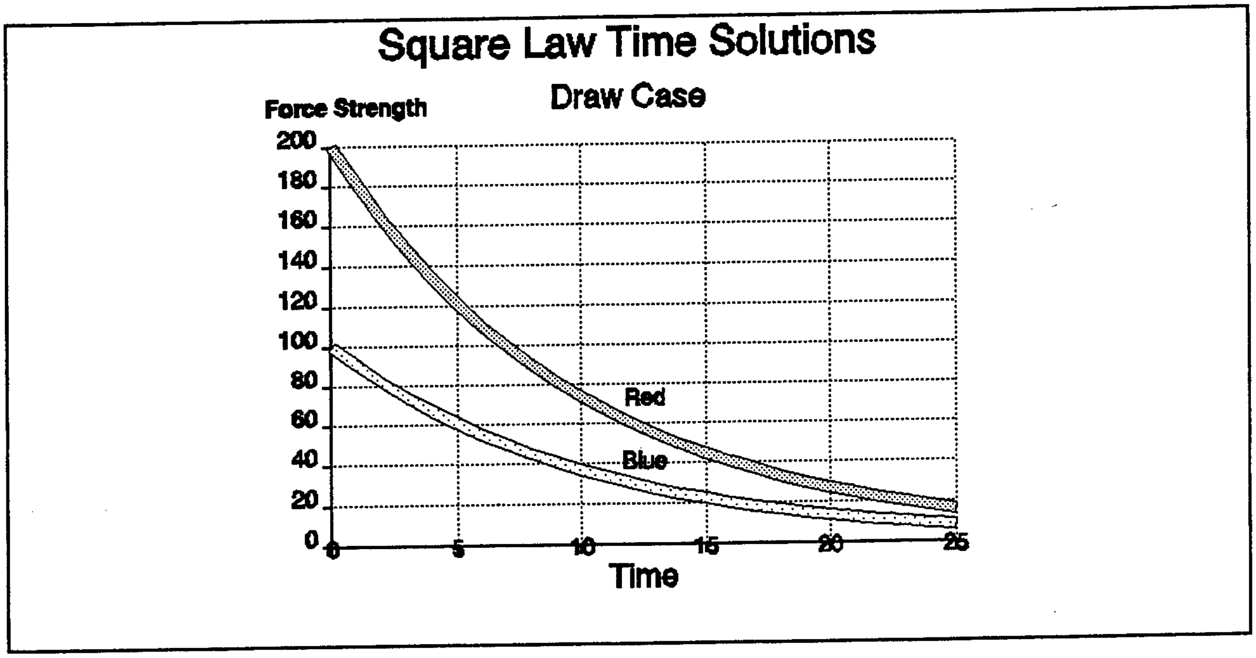


Figure IV.C-5

Figure (IV.C-6) depicts the result if the initial Blue force strength is doubled. This changes to conclusion condition from its zero value for the draw case to a positive value. The rapid attrition of the Red force and the decreased attrition of the Blue force is clearly shown. Halving the initial Blue force strength has the opposite effect, as shown in Figure (IV.C-7).

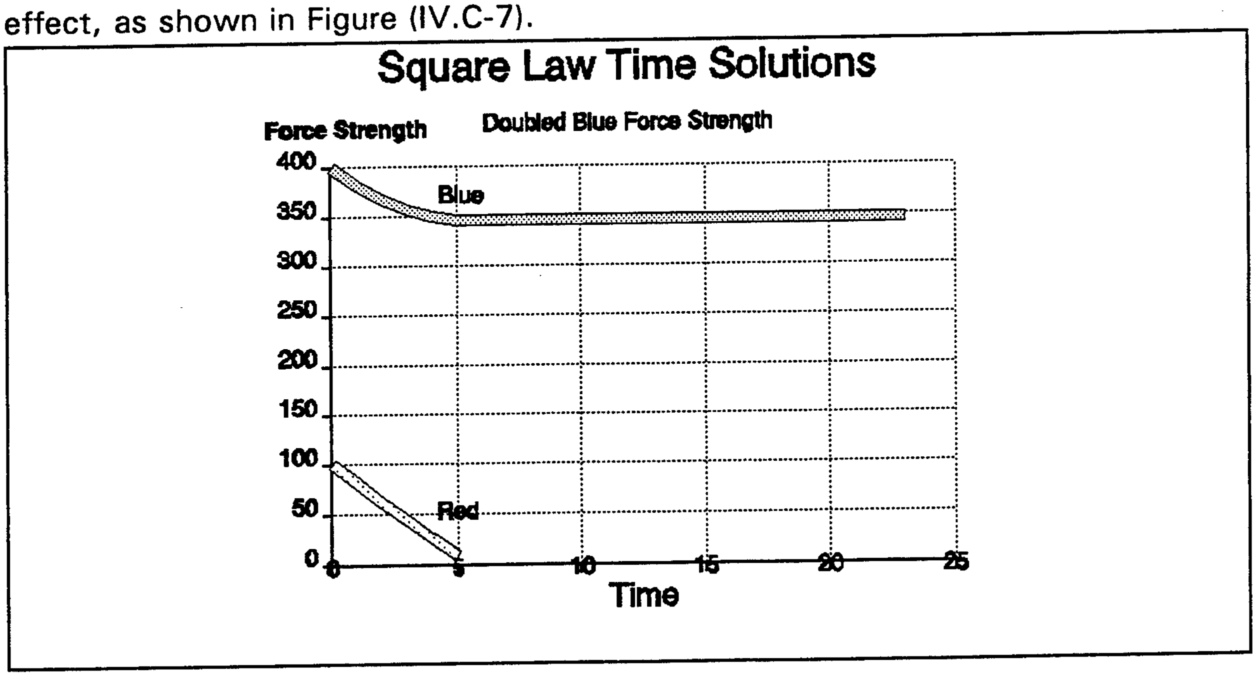


Figure IV.C-6

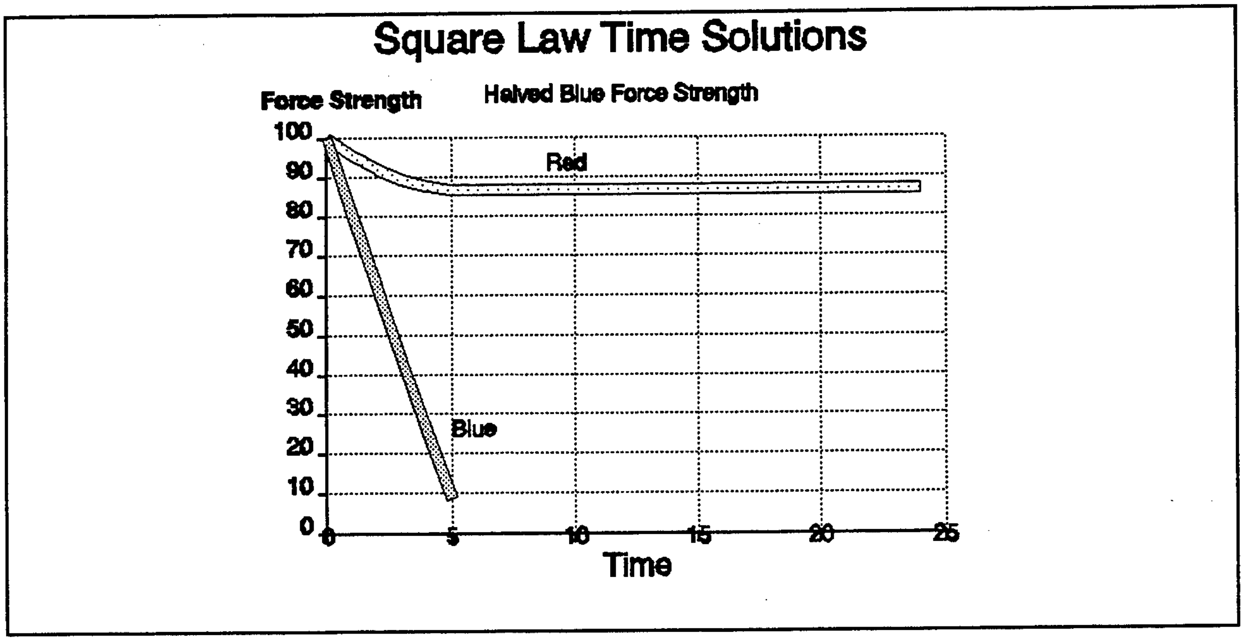


Figure IV.C-7

Changing the attrition rate has a less pronounced effect on the conclusion condition than does changing the initial force strength since the conclusion condition is linear in the attrition rates but quadratic (from whence the name) in the initial force strength. If the Blue force's rate of fire or probability of kill are doubled, the Blue force attrition rate doubles. Comparison with the draw and doubled initial Blue force strength cases, Figures (IV.C-5) and (IV.C-6), respectively, shows the intermediary form of the solutions. Doubling both the rate of fire and the probability of kill of the Blue force has the effect of quadrupling the Blue force attrition rate. In terms of the conclusion condition, this is equivalent to doubling the initial Blue force strength. It also has the effect of doubling the value of γ, so that the attrition process occurs twice as fast as in that of doubling the initial Blue force strength. Comparison with Figure (IV.C-6), the doubled initial Blue force strength shows the same relative losses in both cases: Blue looses 12% of its units in both cases while Red loses 100% of its units, but the attrition process takes half as long in the quadrupled attrition rate case due to the doubling of γ.

IV.D. Lanchester's Mixed Law

The mixed Lanchester law is not explicitly mentioned by Lanchester in **Aircraft in Warfare**; rather, it is suggested by the existence of the linear and quadratic (square) Lanchester laws and the assumptions advanced in the preceding sections. As is the case with the linear and quadratic laws, the mixed law describes combat between two forces. The rate of attrition is driven by the differential equations

$$\frac{dA}{dt} = -\alpha\ B, \qquad \text{(IV.D-1)}$$

which is identical to equation (IV.B.1), and

$$\frac{dB}{dt} = -\beta\ A\ B, \qquad \text{(IV.D-2)}$$

which is identical to equation (IV.A.2). The state solution for these differential equations, derived in Chapter III, is

$$\alpha(B - B_0) = \frac{\beta}{2}(A^2 - A_0^2). \qquad \text{(IV.D-3)}$$

The explicit time solutions of these differential equations are also derived in Chapter III. They differ from the solutions for the linear and quadratic attrition differential equations in that the form of the solution depends on the sign of the quantity

$$\Delta_m \equiv \alpha B_0 - \frac{\beta}{2}A_0^2. \qquad \text{(IV.D-4)}$$

Actually the solutions can be cast into a single functional form if the parameter Δ_m is treated as a complex variable due to the equivalence of the functions tanh(z) and tan(z) for complex argument z. We will not pursue that uniformity here as the mathematics involved are beyond the scope of this book and the resulting functional form does not directly contribute to the discussion of the mixed law.

As in the linear and square law cases, an increase in attrition rate constant/function translates into an increase in force power. For the linear law force (here the Blue force,) an increase in attrition rate constant/function is a direct multiplier of force power while for the quadratic law force (here the Red force,) an increase in attrition rate constant/function directly multiplies force power as the square root of the attrition rate constant/function divided by two. (This factor of two in the denominator must be carried through in the mixed law case because it does not cancel as is the case with the linear law.) This means that if technology is used to increase the attrition rate constant/function, it is more effectively applied to the Blue

force than to the Red force. More exactly, if technology is used to increase Blue's attrition rate constant/function (alpha) by a factor of two, Red can maintain parity only by increasing its attrition rate constant/function by a factor of four.

IV.D.1 Mixed Law Assumptions

The assumptions associated with the mixed law can be carried over directly u from the assumptions associated with the linear and quadratic laws, described previously in Sections IV.B.1 and IV.C.1. The linear law assumptions imply either point attrition (usually direct fire) against targets which are difficult to find or area attrition (usually indirect fire) against a target array whose density changes over time so that the area covered by the target force remains constant. The square law assumptions, on the other hand, imply either point attrition against a target array whose members are easy to find or area attrition against a target array whose density remains constant over time, the area covered by the target force changing over time to keep this density constant. This cross association allows us to describe many types of combat by the three combinations of attrition rate differential equations: linear-linear, quadratic-quadratic, and linear-quadratic (or quadratic-linear.) This association is summarized in Tables IV.D.1 and IV.D.2 which cross correlate the type of fire (direct or area), force disposition (area or density constant), and the density (high or low) to show the type combination of attrition rate differential equations.

In this case, the characteristics direct/area fire, area/density constant, high/low density have been chosen to signify particular aspects of the Lanchester law assumptions. The terminology direct fire is used to signify point attrition while area fire signifies area attrition. Constant density signifies that the force in question maintains a constant areal density of units, thus normally reducing its area of coverage as the number of units decreases through attrition while constant area signifies that the force occupies a constant area during the combat, but that its areal density normally decreases during combat. High density indicates that the units of the force are sufficiently concentrated that target acquisition is fast, while low density signifies that target acquisition is slow, compared to target destruction. This introduction of two different characteristics of unit areal density should, for now, be considered as independent - a force may have a density which is kept constant but which still may be either low or high. Similarly, a force may have a variable density which at any given instant of time may be either high or low. We shall examine these distinctions in density in greater detail in the later chapters of this work which deal with attrition rate constants/functions. The student may also anticipate that we will also deal with some other considerations such as transitions between constant and variable density (constant area occupied,) and the gradations between high and low density which here only serve as limits on whether target acquisition or target destruction processes are dominant.

Some further explanation is also necessary. The student will have noted that the type of fire: area or direct, is the crucial factor on whether constant/variable or high/low density are important in the type of combat being described by one of the three models/laws. What may not have been as obvious is that the attrition differential equation form (linear or quadratic,) for a given force is defined by the density characteristics of that force and the fire type of the opposing force. As an example, the differential equation describing the losses of the red force will be quadratic if the blue force is using direct fire weapons and the red force has high density. The differential equation would be linear if the blue fire were still direct but the red force's density were low.

Examination of this table reveals that inclusion of the mixed law permits the modeling of combat between forces in a manner which the strict linear and quadratic laws would not permit. Specifically, we see that the linear law would allow consideration of the following forms of combat:

Table IV.D.1

Red	Blue
Area Fire, Density Constant	Area Fire, Density Constant
Direct Fire, Low Density	Direct Fire, Low Density
Area Fire, Low Density	Direct Fire, Density Constant
Direct Fire, Density Constant	Area Fire, Low Density

This short table illustrates the cross relationship between fire type for one force and the density characteristics of the other force. The same table for the quadratic law is:

Table IV.D.2

Red	Blue
Area Fire, Area Constant	Area Fire, Area Constant
Direct Fire, High Density	Direct Fire, High Density
Area Fire, High Density	Direct Fire, Area Constant
Direct Fire, Area Constant	Area Fire, High Density

It is readily obvious that out of 16 possible combinations of fire type and density characteristics (4 per force,) that the original linear and quadratic Lanchester laws will only admit to modeling 8 combinations. The rest of the possible combinations fall under the mixed law. (Note that these 16 combinations are not exhaustive - they merely cover the extremes permitted under the basic assumptions associated with the Lanchester laws.)

IV.D.2 Mixed Law State Solution

As in the previous cases, the quantity Δ_m is a measure of the forces remaining if the combat is carried to a conclusion. Unlike the Δ's defined in the linear and quadratic law cases, this Δ is not symmetric in the force strengths. Thus, a small change in the red force strength (we have explicitly assumed that the red force is linear-like while the blue force is quadratic-like - this can be reversed with only the necessary symmetric swapping of force strengths and attrition rates,) will have a much greater effect on the value of Δm than will an equal change in the blue force strength. While we might normally expect α to be much larger than β to correct for this, we must note that the attrition rates are constants (or functions,) and therefore only point values. Thus, in mixed combat, there is a great advantage to the linear-like force in greater numbers if the combat were to be carried to a conclusion. This can readily be seen in Figure (IV.D-1) where we plot blue force strength versus red force strength for two value of Δm which are equal in magnitude but opposite in sign. The draw case, unlike the other two state solutions, is not a straight line, but rather is a parabola. This form is the direct result of the asymmetric nature of the state solution. The curvature of the graph is readily apparent. However, just as the state solutions for the linear and quadratic state solutions are symmetric about the draw case for opposite values of (delta), so too are the solutions for the mixed law. This symmetry is somewhat more difficult to see due to the curvature of the draw case. If the student can imagine transforming the draw case state solution to a straight line, and mentally repeat these operations on the two other state solutions in the figure, then the symmetric arrangement can be visualized.

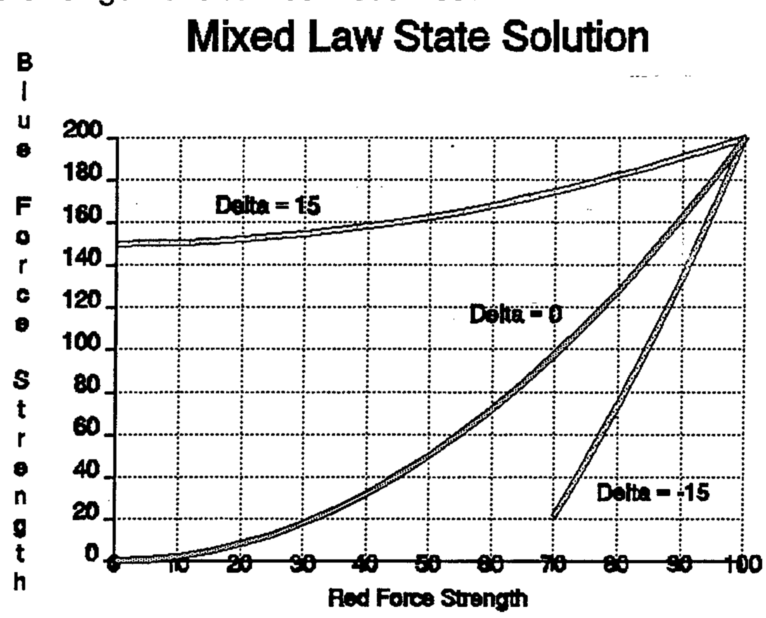


Figure IV.D-1

Note that changes in the value of Δ_m cause different values in the remaining force strength at conclusion. Of course, the draw case results in zero force strength on both sides. Thus each side takes 100% losses. For a Δ_m value of 15 however, the red force takes 100% losses, while the blue force takes 25% losses. Alternately, for a Δ_m value of -15, the blue force takes 100% losses, while the red force takes about 35% losses. This asymmetry is the direct result of the values of the attrition rates , and illustrates the effect only of changing the attrition rates, not the initial force strengths.

IV.D.3 Mixed Law Time Solution

The time solution of the Lanchester mixed attrition differential equation for red force linear-like, and blue force quadratic-like, derived in Chapter III, are

$$A(t) = \frac{A_0 - \sqrt{\frac{2|\Delta_m|}{\beta}}\tan(\eta t)}{1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\tan(\eta t)}, \Delta_m > 0$$

$$B(t) = \frac{B_0 \sec(\eta t)^2}{\left[1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\tan(\eta t)\right]^2}, \Delta_m > 0 \qquad \text{(IV.D-5)}$$

and

$$A(t) = \frac{A_0 + \sqrt{\frac{2|\Delta_m|}{\beta}}\tanh(\eta t)}{1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\tanh(\eta t)}, \Delta_m < 0$$

$$B(t) = \frac{B_0\, sech(\eta t)^2}{\left[1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\tanh(\eta t)\right]^2}, \Delta_m < 0 \qquad \text{(IV.D-6)}$$

where:

$$\eta = \sqrt{\frac{\beta\,|\Delta_m|}{2}}. \qquad \text{(IV.D-7)}$$

We note immediately that there are two forms of these solutions which depend on the sign of Δ_m. We can directly reduce these solutions for the draw case if we note that $\tanh(x) = \tan(x) \to x$ and $\sec(x) = \mathrm{sech}(x) \to 1$ as $x \to 0$ (to first order in x). This allows us to write

$$A(t) = \frac{A_0 + \sqrt{\frac{2|\Delta_m|}{\beta}}\,(\eta t)}{1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\,(\eta t)}, \ \Delta_m < 0$$

$$B(t) = B_0 \frac{}{\left[1 + A_0\sqrt{\frac{\beta}{2|\Delta_m|}}\,(\eta t)\right]^2}, \ \Delta_m < 0 \qquad \textbf{(IV.D-8)}$$

From the definition of η, these equations may be rewritten as

$$A(t) = \frac{A_0 - |\Delta_m| t}{1 + A_0 \beta t}, \ \Delta_M < 0$$

$$B(t) = \frac{B_0}{[1 + A_0 \beta t]^2}, \ \Delta_m < 0. \qquad \textbf{(IV.D-9)}$$

Since the draw case means that $\Delta = 0$, we must rewrite Equation (IV.D-5a) or (IV.D-6a) as

$$A(t) = \frac{A_0}{1 + A_0 \beta t}, \qquad \textbf{(IV.D-10)}$$

while Equation (IV.D-5b) or (IV.D-6b) may be used without alteration since it does not explicitly contain Δ_m.

The explicit time solutions for these two values of Δ_m are shown in Figures (IV.D-2) and (IV.DC-3) for the Red and Blue force strengths, respectively. As in the previous figures presented in this chapter, the units of the time variable are chosen arbitrarily. It may be seen that the positive Δm solution reaches zero faster (Red force - quadratic-like) than the negative (delta)m solution (Blue force - linear-like) does.

Another way to vary the value of Δm is to alter the values of α and β. The changes of the solutions for variations of $\pm 50\%$ in the value of β are shown in Figures (IV.D-4) and (IV.D-5) for the Red and Blue force strengths, respectively. These variations are executed relative to the case of $\Delta_m = 15$. As noted earlier, the effect of the variation in β scales as a square root change in A_0. Although it is difficult to see, the effect of these variations on the time solutions are approximately equal in relative magnitude for these effective small variations in the solutions.

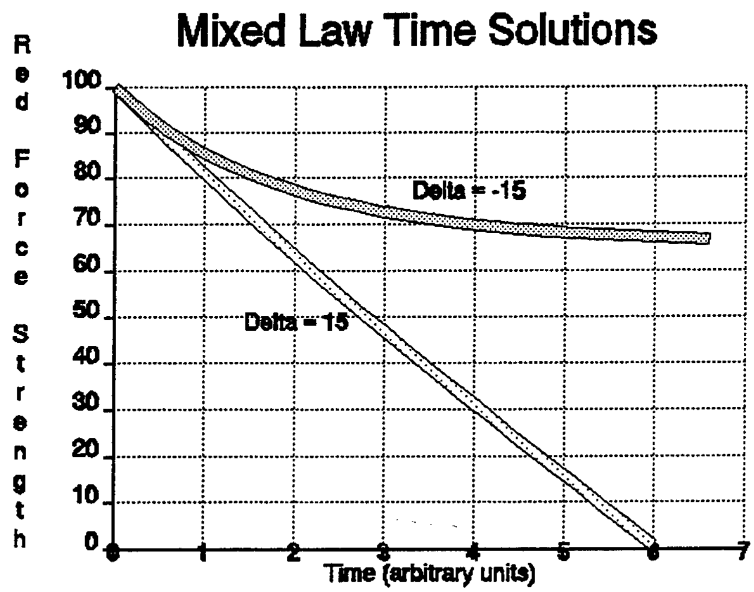


Figure IV.D-2

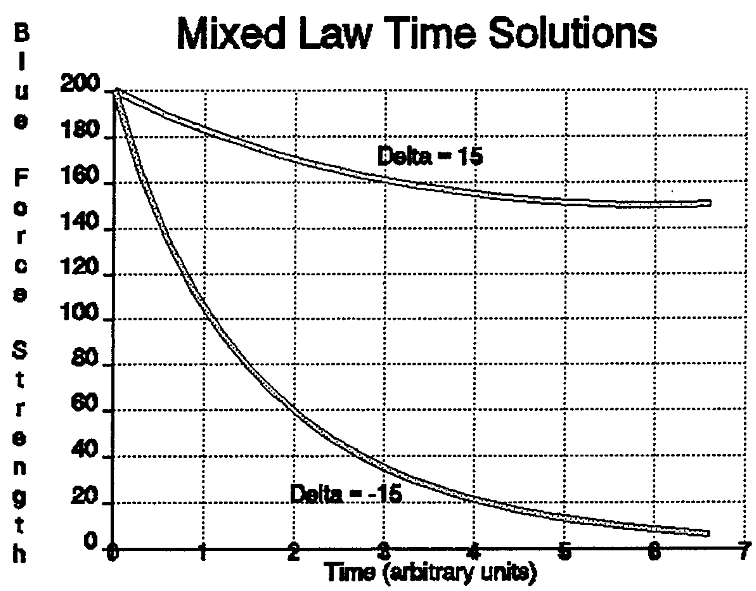


Figure IV.D-3

Variations in α of similar magnitude ($\pm 50\%$) are shown in Figures (IV.D-6) and (IV.D-7) for the Red and Blue force strengths, respectively. Note the relatively greater changes in the shapes of the solutions. This is the result of relatively greater changes in the values of Δm about the base case (15) value. The changes in the Blue force strength solutions are actually less than those due to the variations in β; the changes in the Red force strength solutions are decidedly pronounced.

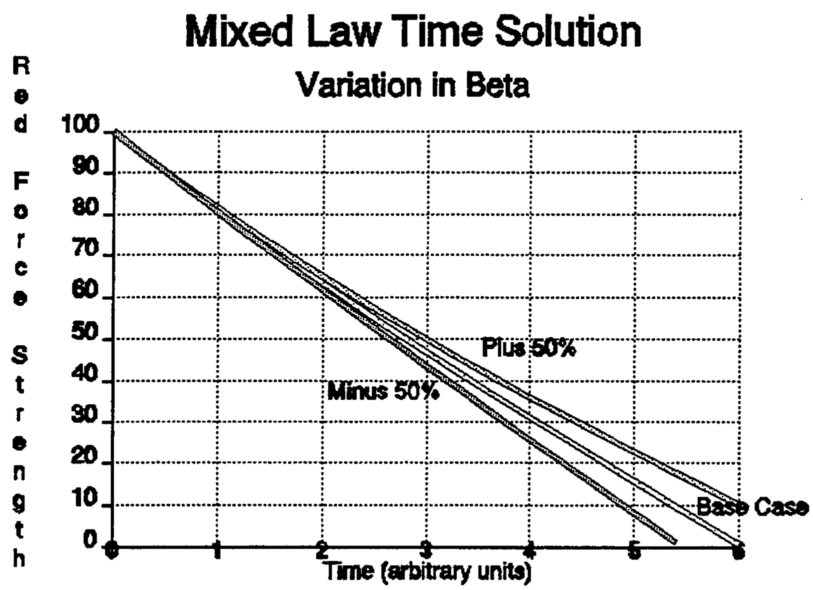


Figure IV.D-4

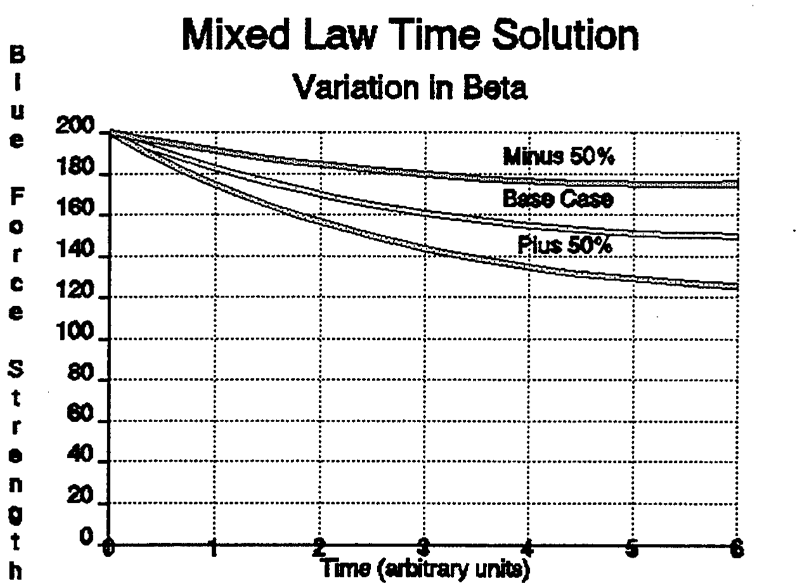


Figure IV.D-5

The other way to vary the solutions is to alter the value of the initial force strengths. Variations in the value of the initial Red force strength (B_0) of $\pm 50\%$ are shown in Figures (IV.D-8) and (IV.D-9). As expected, the Red force strength time solutions are essentially parallel. Variations in the value of the initial Blue force strength (A_0) of $\pm 50\%$ are shown in Figures (IV.D-10) and (IV.D-11). Again, the Blue

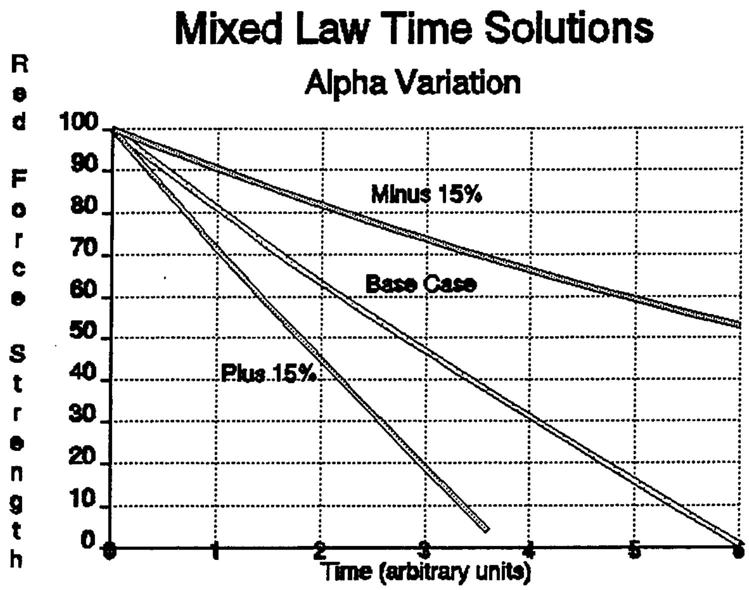


Figure IV.D-6

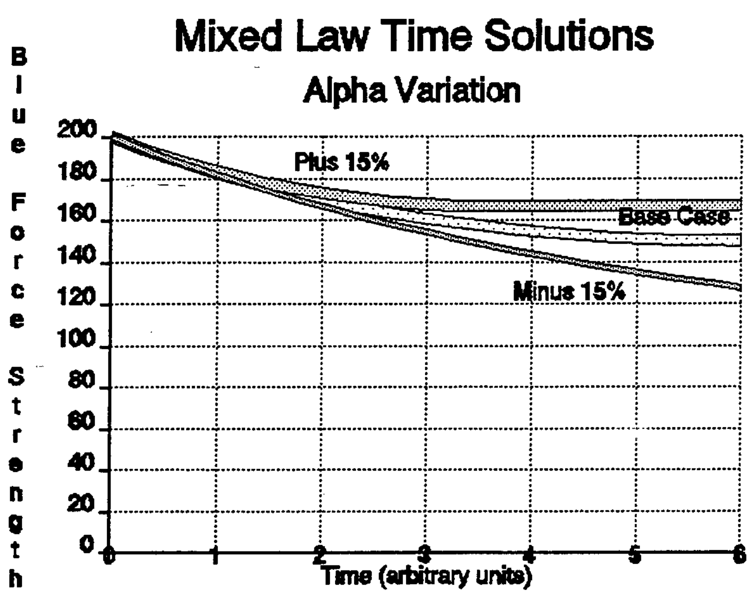


Figure IV.D-7

force strength time solutions are essentially parallel. We see here effects much as predicted from the form of the state solution. Increasing the initial strength of the Red force prolongs the duration of the battle (if carried to a conclusion or a percentage loss), while increasing the initial strength of the Blue force shortens the battle (under the same conditions.) For the cases studied here (admittedly for positive Δ_m,) indicate that increases in initial force strength tend to favor the linear-like force more than the

square-like force while increases in the quadratic-like attrition rate (α) tend to favor the square-like force more than the linear-like force (β).

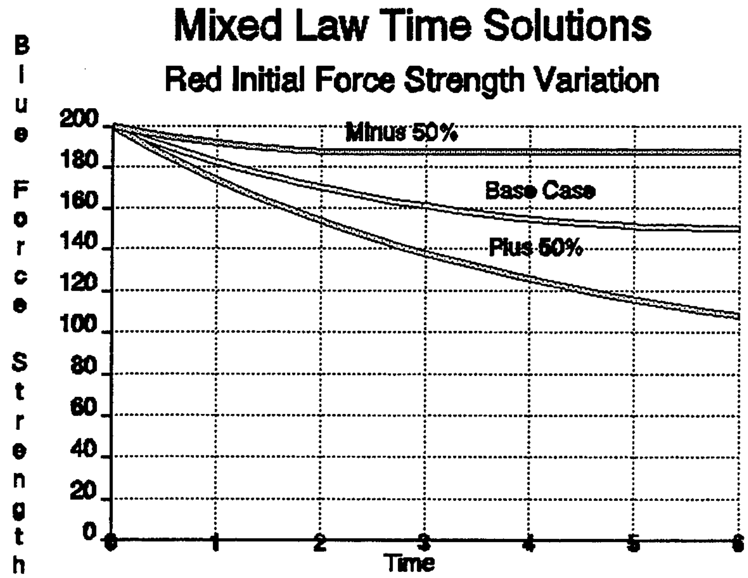


Figure IV.D-8

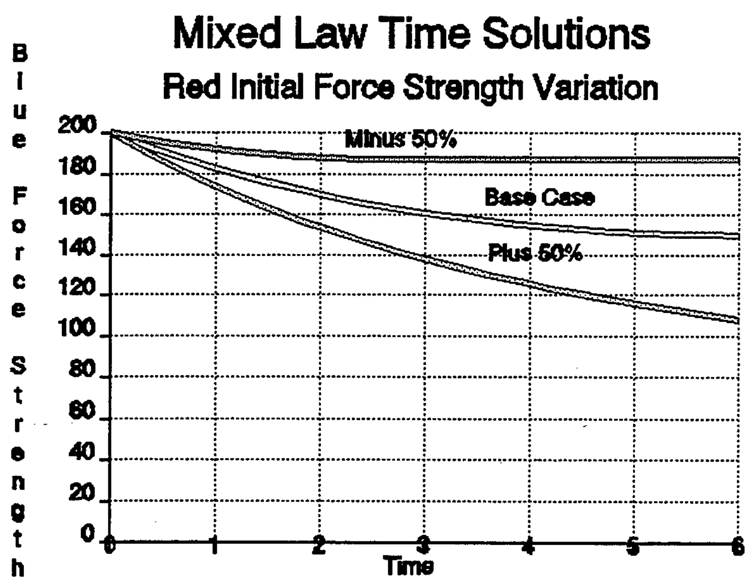


Figure IV.D-9

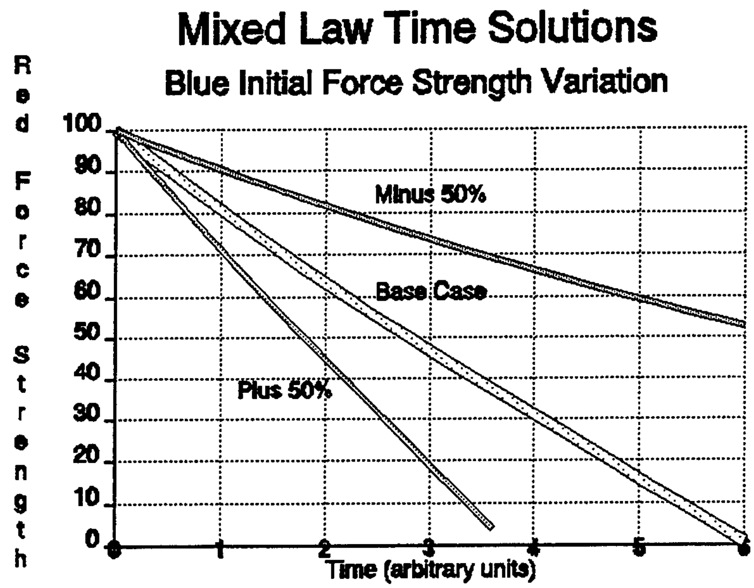


Figure IV.D-10

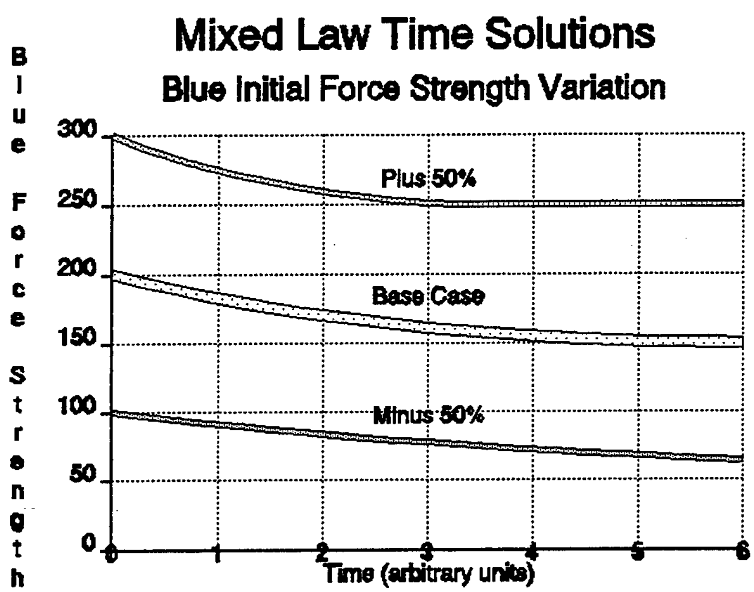


Figure IV.D-11

IV.E Force Ratio

The time solution for the force ratio were derived in Chapter III. As noted in that chapter, the force ratio, defined by

$$\rho(t) = \frac{\rho_0 - \delta \tanh(\gamma t)}{1 - \frac{\rho_0}{\delta} \tanh(\gamma t)}, \qquad \text{(IV.E-1)}$$

where γ and δ were defined in Section IV.C., can only be defined in closed form for the quadratic Lanchester law. The force ratio can be calculated for the linear and mixed Lanchester laws, but only from the explicit time solutions (or from one of the pair of time solution equations and the state solution,) but the force ratio for these laws cannot be explicitly defined only in terms of the initial force ratio and the attrition rate constants. This can only be done for the quadratic Lanchester law.

The cases of initial force strength and attrition rate variations presented in Section IV.C are reproduced in Figures (IV.E-1) and (IV.E-2), respectively. The draw case clearly shows the constancy of the force ratio. This follows from the fact that the initial force ratio for the draw case is exactly equal to δ. (The student can easily confirm this for himself - we shall explicitly derive this result in the next chapter.) Since changing the initial Blue force strength completely shifts the sign of Δ_2 to be positive (doubling B_0) or negative (halving B_0,) the variation of the value of ρ on $[0, \infty)$ is clearly indicated in Figure (IV.E-1), although we have truncated the halved Blue solution shy of conclusion to avoid warping the figure excessively. The effect of increasing the Blue attrition rate parameters (Rate Of Fire and Probability of Kill - using the simple attrition model developed in Section IV.C,) shows the same behavior that we had earlier noted in that section - namely that changes in the attrition rate have an effect which scales as the square of changes in initial force strength.

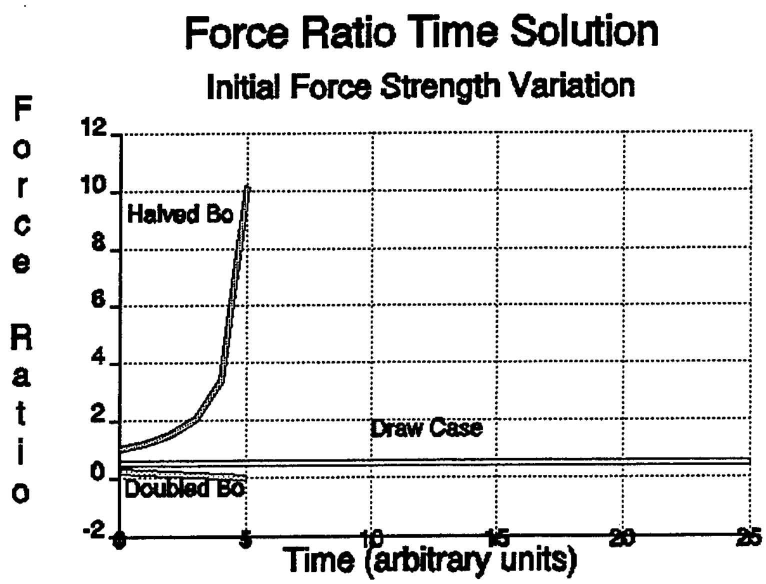


Figure IV.D-1

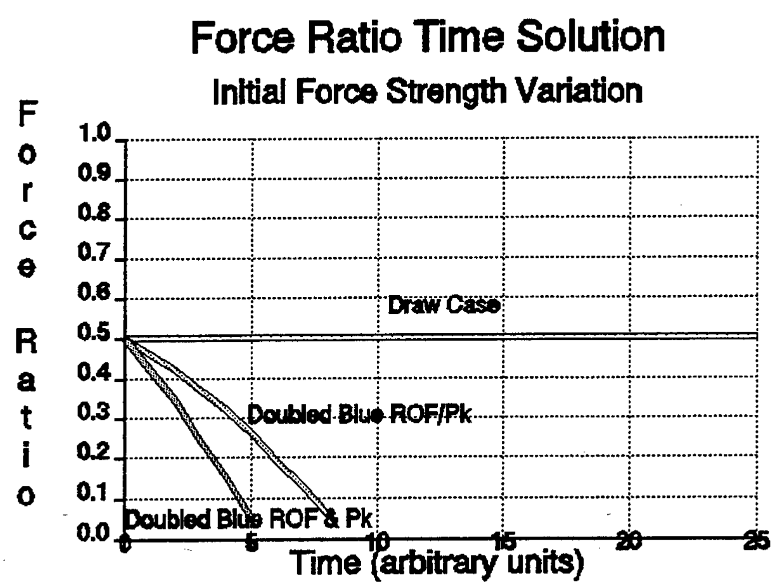


Figure IV.D-2

IV.F Final Note

In the next chapter, we will present the Ironman analyses which provide further insight into the nature of the attrition rate constants/functions. This chapter builds on the mathematical tools which were built in Chapter III, and the analyses and

assumptions laid down in this chapter.

In Chapter VI, we will derive several additional mathematics tools and equations which are useful in the study of Lanchestrian attrition mechanics. These will include formal derivations of the draw case attrition time solutions which have been sketched in this chapter as *ad hoc*. The remainder of the expositions to be presented in that chapter will largely deal with approximations necessary for consideration of more complicated attrition differential equations than the pure Lanchestrian forms described in this and the last chapter.

V. THE IRONMAN ANALYSES

V.A. Introduction

This chapter is devoted to what I call the Ironman Analyses. At times, I have used the term Ironman Theorems because these analyses can be cast in the form of theorems. However, with the time and cogitation taken in developing these analyses, I find that they lack the rigor to honestly be called theorems and that terming them analyses is more useful.

The Ironman Analyses are a tool for understanding the nature of the Lanchester Equations and their accompanying attrition rates. As we shall see, the Lanchester Equations and the attrition rates are a dualism. The one defines the other. The analyses bring the interpretation that the Lanchester Differential Equations are actually the definitions of the attrition rates.

The motivation for the Ironman Analyses arises from Bonder's work. In his thesis, he provides what is called here Bonder's Equations,

$$\alpha = T^{-1}. \qquad \text{(V.A-1)}$$

where α is the attrition rate and T is the expected time for one unit to kill another (enemy) unit. Bonder presents this equation without adequate analytical underpinnings. Despite its apparent and intuitive correctness, some additional basis for this equation seemed to be needed. Thus, the Ironman Analyses.

Central to these analyses is the concept of the Ironman. In simplest terms, an Ironman is a foe who cannot be destroyed or attrited. As such, a force comprised of Ironmen cannot change with time. That force is a constant. The attrition rate acting on it is zero. As a result, the pair of Lanchester Differential Equations effectively reduces to one equation which has a simple solution. This solution provides a direct definition of the attrition rate acting on the Ironman force's foe.

The Ironman Analyses represent a special, restrictive case of the Lanchester Differential Equations or, perhaps more generally, of the transport theory of warfare. As a class, they are comparable to the gedanken or thought experiments of Quantum Mechanics. They provide a similar function in providing insight into the processes of attrition just as the thought experiments of Quantum Mechanics provide insight into its workings.

We shall be concerned with two types of Ironman Analyses in this chapter: deterministic and statistical. The deterministic analyses are straightforward from the basic Lanchester Differential Equations. The statistical analyses are less

straightforward, but yield greater insight into the interplay of possible statistical forces with attrition. Further, they lay a groundwork for developing the statistical forms of Lanchester's Equations in a subsequent chapter.

V.B. Deterministic Linear Ironman

Since the Ironman cannot be attrited, the Linear Lanchester Differential Equations, Equations (III.A-2) and (III.A-3) reduce to

$$\begin{aligned} \frac{dA}{dt} &= -\alpha\ A\ B_0, \\ B(t) &= B_0, \\ \beta &= 0, \end{aligned} \tag{V.B-1}$$

if we take the B force to be comprised of Ironmen. For convenience, we shall restrict the B force to have a strength of one. As a result, Equation (V.B-1) can be solved directly as

$$A(t) = A_0\ e^{-\alpha t} \tag{V.B-2}$$

which we will also write in the form

$$\alpha\ t = \ln\left(\frac{A_0}{A(t)}\right). \tag{V.B-3}$$

If we now take this equation for a particular characteristic time τ to be the time to kill one enemy unit, then

$$\alpha\ \tau = \ln\left(\frac{A_0}{A_0 - 1}\right). \tag{V.B-4}$$

This may be rewritten as

$$\alpha\ \tau = -\ln\left(1 - \frac{1}{A_0}\right), \tag{V.B-5}$$

which we may expand to first order since $A_0 \gg 1$ by assumption (using the logarithmic expansion formula.) This gives,

$$\alpha\ \tau \simeq A_0^{-1}, \tag{V.B-6}$$

or

$$\alpha \approx (\tau\, A_0)^{-1}. \qquad \text{(V.B-7)}$$

As an example, let us now consider a combat unit that requires some amount of time t_1 to find the target and then fires on the target at intervals t_2. Let us further stipulate that every time the target takes fire, it loses a fraction f of its strength. The strength of the force as a function of time may be written immediately as

$$A(t_1 + nt_2) = (1 - f)^n A_0, \qquad \text{(V.B-8)}$$

after the force has been fired on n times. (For the sake of maintaining the discussion, we shall not rigorously require that n be treated as an integer.)

If we now relax the requirement, for the moment, that n be an integer, we may calculate a value for the (probably less than 1) number of fires necessary to kill on target, n^*. Obviously

$$\tau = t_1 + n^* t_2, \qquad \text{(V.B-9)}$$

and n^* may be obtained from

$$(1 - f)^{n^*} A_0 = A_0 - 1, \qquad \text{(V.B-10)}$$

if we assume $f \ll 1$, as

$$n^* f A_0 \approx 1. \qquad \text{(V.B-11)}$$

Thus,

$$\tau = t_1 + \frac{t_2}{f A_0}, \qquad \text{(V.B-12)}$$

and the attrition rate is just

$$\alpha \approx \frac{f}{A_0 f t_1 + t_2}. \qquad \text{(V.B-13)}$$

If we cannot make the assumption that $f \ll 1$, then equation (V.B-10) may be rearranged, the logarithm taken, to yield,

$$n^* = \frac{\ln\left(1 - \frac{1}{A_0}\right)}{\ln(1 - f)} \tag{V.B-14}$$
$$\approx -\frac{1}{A_0 \ln(1 - f)},\ A_0 > 1,$$

which gives an attrition rate of

$$\alpha = \frac{A_0^{-1} \ln(1 - f)}{t_1 \ln(1 - f) + t_2 \ln\left(1 - \frac{1}{A_0}\right)} \tag{V.B-15}$$
$$\approx \frac{\ln(1 - f)}{A_0\, t_1 \ln(1 - f) - t_2},\ A_0 > 1.$$

Note that since $\ln(1 - f) < 1$, the minus sign in front of the t_2 does not decrease the denominator.

We shall examine the consequences of the linear attrition rate further in the analysis of the statistical forms of the linear Lanchester Equations.

V.C. Deterministic Quadratic Lanchester Equations

Under the assumptions of the Ironman Analysis, the Quadratic Lanchester Differential Equations develop in an identical manner as the linear equations. They become,

$$\frac{dA}{dt} = -\alpha\ B, \qquad B(t) = B_0, \qquad \beta = 0. \tag{V.C-1}$$

For convenience, we again take the Ironman force to have a strength of one. The solution to the differential equation is then

$$A(t) = A_0 - \alpha\ t, \tag{V.C-2}$$

and if we again adopt a characteristic time, the quadratic attrition rate takes the form

$$\alpha = \frac{A_0 - A(\tau)}{\tau}. \qquad \text{(V.C-3)}$$

In this quadratic case, the meaning of the characteristic time seems clear if we interpret it as the time required to attrit the A force's strength by one. That is

$$A(\tau) = A_0 - 1. \qquad \text{(V.C-4)}$$

If we adopt a simple model of a combat unit that takes time t_1 to find the target and thereafter fires at intervals t_2 with a (constant) single shot kill probability p, the expected number of shots fired to kill the target is 1/p. The quadratic attrition rate for this unit is

$$\alpha = \frac{1}{t_1 + \frac{t_2}{p}}. \qquad \text{(V.C-5)}$$

V.D. Comment on the Combat Unit

The type of unit described above obeys what is known as a geometric probability distribution. This distribution describes a sequence of (presumably identical) trials (shots) which may be infinite in number. The distribution states that the probability of the event (in our case, a kill,) occurring on the n^{th} trial is $(1 - p)^{n-1}$ p where p is the probability that the event will occur on the first (or any other) trial.

The form of the distribution can be directly seen if one considers that the probability of the event not occurring (commonly called a failure) on the first trial is (1 - p). The probability of the event occurring (commonly called a success) on the second trial is the probability of the first trial being a failure [(1 - p)] times the independent probability of the second event being a success [p]. The conditional (or total) probability of the second event being a success is then (1 - p) p.

Similarly, the total probability of success on the third trial is the probability of the first and second trials being failures [(1 - p)(1 - p) = $(1 - p)^2$] times the conditional probability of success on the third trial [p] or $(1 - p)^3$ p.

As a model of a combat unit, this is exceedingly simple. The model requires that the probability of kill of the unit against its target be the same for each shot. Despite its simplicity, however, this model is of increasing validity as weapon systems improve. Thus, for modern weapon systems, this model may be valid, while for older weapon systems it may likely not be.

If p is the probability of kill per shot, and we designate the quantity (1 - p) by the variable q, then the total probability of kill after N shots is

$$P(N) = \sum_{i=1}^{N} q^{i-1}\, p. \tag{V.D-1}$$

If we allow an infinite number of shots to be fired (or trials to be made,) the total probability of kill is

$$P(\infty) = p \sum_{i=1}^{\infty} q^{i-1}, \tag{V.D-2}$$

which may be rewritten as

$$P(\infty) = p \sum_{i=0}^{\infty} q^{i}, \tag{V.D-3}$$

by change of index. This sum is exactly summable using,

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^{i},\ |x| < 1, \tag{V.D-4}$$

which is called a geometric series. Note that in our case, q < 1 always. An interesting corollary is the case where N shots are fired, that is

$$\sum_{i=0}^{N-1} x^{i} = \frac{1-x^{N}}{1-x}. \tag{V.D-5}$$

From Equation (V.D-4), we may immediately see that Equation (V.D-3) sums to

$$P(\infty) = \frac{p}{1-q} = \frac{p}{p} = 1, \tag{V.D-6}$$

so that the distribution is normalized.

We may calculate the time to kill the target (the expected time to kill) using this distribution. Using

$$\tau = t_1 + i\, t_2, \tag{V.D-7}$$

where i is the number of shots fired, we may form the summation

$$\begin{aligned} <\tau> &= p \sum_{i=1}^{\infty} q^{i-1} (t_1 + i\, t_2) \\ &= t_1 + t_2\, p \sum_{i=0}^{\infty} q^{i} (i + 1), \end{aligned} \tag{V.D-8}$$

where the factor t_1 comes through directly because it is deterministic, and $<>$ indicates an ensemble average (expected value.) (In practice, the performance of the unit may have more than one probability distribution associated with it. For example, finding the target and killing the target may be represented by two different probability distributions. When more than one probability distribution is present, they are each treated in much the same manner as above.) We may evaluate the summation by noting that

$$\frac{dq^{i+1}}{dq} = (i + 1)\, q^{i}. \tag{V.D-9}$$

This gives

$$<\tau> = t_1 + t_2\, p \sum_{i=1}^{\infty} \frac{dq^{i+1}}{dq}. \tag{V.D-10}$$

If we assume the summation and derivative to be mutually independent, then we may swap their order, and rewrite this as

$$\begin{aligned} <\tau> &= t_1 + t_2\, p \frac{d}{dq} \sum_{i=0} q^{i} \\ &= t_1 + t_2\, p \frac{d}{dq} \frac{1}{1 - q} \\ &= t_1 + t_2\, p \frac{1}{(1 - q)^2} \\ &= t_1 + t_2 \frac{p}{p^2} \\ &= t_1 + \frac{t_2}{p}, \end{aligned} \tag{V.D-11}$$

which is the result claimed in Equation (V.C-5).

Note that this model becomes invalid if the firer is allowed only a finite number of shots (e.g. limited ammunition,) since either a rearm time must be introduced or a probability of the target surviving. In the latter case, the expected value of the time to kill becomes infinite.

Another distribution of interest is the Hypergeometric probability distribution. This distribution deals with a finite population of items (say N) which are of two types (say red and blue in color.) The quantity of interest described by this distribution is the number of items of a given type (color) that are selected given some total number of trials of selection.

For example, if there are n red items (and thereby N - n blue items) and M items are selected (M trials,) then the probability that m red items will be selected is

$$p(m) = \frac{\binom{n}{m}\binom{N-n}{M-m}}{\binom{N}{M}}, \tag{V.D-12}$$

while the expectation value associated with this distribution is

$$<m> = \frac{n\,M}{N} \tag{V.D-13}$$

Yet another useful distribution is the Poisson. This distribution is used to describe the likelihood of a number of identical events occurring over a measurable interval (of time or distance usually.) Specifically, the probability of one event occurring over an interval dx is λ dx. The probability of the event not occurring is (1 - λ dx). By this definition, two or more events may not occur in the same interval dx. (On another note, we could say that the probability of n events occurring in interval dx is λ^n dx^n. The probability of any events occurring is

$$P(n > 0) = \sum_{n=1}^{\infty} (\lambda\; dx)^n, \tag{V.D-14}$$

which we may immediately sum using the geometric series and some algebra as

$$P(n > 0) = \frac{\lambda\; dx}{1 - \lambda\; dx}. \tag{V.D-15}$$

The probability of no events occurring is one minus this quantity which we may expand as

$$p(n > 0) \approx 1 - \lambda\; dx - \lambda^2\; dx^2 - HOT. \tag{V.D-16}$$

As an approximation then, we may see that we must take

$$\lambda^2\, dx^2 < 1, \tag{V.D-17}$$

to satisfactorily ignore multiple simultaneous events.)

Returning now to our exposition on the Poisson probability distribution, the probability of n events occurring in the interval (0,x) is then

$$P_n(x) = \frac{(\lambda x)^n}{n!}\, e^{-\lambda x}, \tag{V.D-18}$$

and the expected number of events in the interval is

$$<n>(x) = \lambda\, x. \tag{V.D-19}$$

To put this in context, if a combat unit fires every Δt seconds and has a probability of kill p, then the expected number of kills over time is

$$<kills>(t) = \frac{p\, t}{\Delta t}, \tag{V.D-20}$$

since $\lambda = p/\Delta t$, and the probability of n kills is just

$$P_{n\ kills} = \left[\frac{p\ t}{\Delta t}\right]^n \frac{e^{-\frac{p\,t}{\Delta t}}}{n!}. \tag{V.D-21}$$

If we use our previous example from the geometric probability distribution, then the expected number of kills in time $<\tau>$ is one, and the expected number of kills over an interval t is

$$<kills> = \frac{t}{<\tau>}, \tag{V.D-22}$$

and the probability of n kills in time t is

$$P_{n\ kills} = \left[\frac{t}{<\tau>}\right]^n \frac{e^{-\frac{t}{<\tau>}}}{n!}. \tag{V.D-23}$$

V.E. Statistical Linear Lanchester Equation

Let us consider the case where force A has an initial strength A_0 spread over an area L and that force B fires every time interval Δt. Every fire has a effective area d with a probability of kill p. (We obviously neglect some delivery accuracy here.) After the first fire, A has taken loses

$$\frac{A_0\, p\, d}{L} = A_0 f; \left(f = \frac{p\, d}{L} \right), \tag{V.E-1}$$

which is a fractional loss f. If we designate the strength of the A force after n fires as A_n, then the strength after one fire is

$$A_1 = A_0 \,(1 - f). \tag{V.E-2}$$

Two possible extremes may be considered here for the mobility and reactiveness of the A force: either force A can move rapidly enough to reposition itself ($v_m\, \Delta t \gg \sqrt{d}$) where v_m = movement rate of force A) or it cannot ($v_m\, \Delta t \ll \sqrt{d}$.) In the first case, the areal density of A, A_n/L is a constant over L. In the second case, the areal density is not constant over L. We shall consider each of these cases in turn.

The first case, where repositioning of forces is accomplished, is the simpler of the two cases. We may even extend the analysis to incorporate multiple firers in the B force.

After n fires, the strength of the A force is either

$$A_n = A_0 \,(1 - f)^{Bn}, \tag{V.E-3}$$

if no two firers in the B force fire at the same point at the same time, or

$$A_n = A_0 \,(1 - B f)^{n}, \tag{V.E-4}$$

if all of the firers in the B force fire at the same point at the same time.

Regardless of which extreme of B force firing doctrine we select, we require that

$$f \ll 1, \tag{V.E-5}$$

and

$$B f < 1. \tag{V.E-6}$$

Not only are the individual fires of B not very lethal, but the total salvo fires of B are not very lethal. (Obviously this does not apply to weapons of mass destruction - nuclear, chemical or biological weapons, and the assumption of low lethality for salvo or volley fire is not necessarily valid.[1])

If we now relax the requirement that n be an integer (make the transition from impulsive to continuous attrition,) and replace it with the variable t/Δt, then Equations (V.E-3) and (V.E-4) may be rewritten as

$$A9t) = A_0 \, (1 - f)^{\frac{Bt}{\Delta t}}, \tag{V.E-7}$$

and

$$A(t) = A_0 \, (1 - B f)^{\frac{t}{\Delta t}}. \tag{V.E-8}$$

It is convenient to rewrite Equation (V.E-8) as

$$A(t) = A_0 \, e^{\frac{t}{\Delta t} \ln(1 - Bf)}, \tag{V.E-9}$$

and since B f ≪ 1, we may expand Equation (V.E-9) using

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} \; .... \tag{V.E-10}$$

and retain only the first order term, thus giving

$$A(t) = A_0 \, e^{-\frac{Bft}{\Delta t}}. \tag{V.E-11}$$

Note that the use of Equation (V.E-10) with Equation (V.E-7) and the condition that f ≪ 1, gives the same result. Thus, as long as f ≪ 1 and B f ≪ 1, the distribution of force B's fires is not important (to first order in f!)

If we now differentiate Equation (V.E-11) with respect to time, we get

$$\frac{dA}{dt} = -\frac{f}{\Delta t} B A, \tag{V.E-12}$$

which is the Linear Lanchester Differential Equation with an attrition rate of

$$\begin{aligned}\alpha &= \frac{f}{\Delta t} \\ &= \frac{p\,d}{L\,\Delta t}.\end{aligned} \tag{V.E-13}$$

The attrition rate of the linear case is thus the ratio of the area of effect of the fire to the total area covered by the A force (d/L) divided by the expected time to kill for a fire ($\Delta t/p$). If we return to Equation (V.B-10), neglect t_1, and expand the logarithm with Equation (V.D-10) using the same restriction, we get

$$\alpha \approx \frac{f}{t_2}, \tag{V.E-14}$$

which agrees with Equation (V.E-13).

When the fire is too rapid for the A force to relocate, a more complex mathematical problem occurs. To facilitate our analysis, we shall restrict the B force so that all fires are at different places at the same time, and that several fires are repeated over time at the same aim point. We shall further assume that no difference need be drawn between those of the A force which have been killed as the result of a fire and those which have not, at least in terms of determining subsequent kills. That is, each subsequent fire 'kills' a fraction f of the A force in the area of effect (the kill area,) but some of that fraction may already have been killed. As an illustration, take f to be 1%. The first fire kills 1% of the A force in the kill area. The second fire also kills 1% of the A force, but 1% of that 1% were already killed in the first fire. Thus the total fraction killed after the second fire is 1.99%, not 2.00%.

One way of treating this reduction in the effective fraction killed is to note that the successive fraction killed by each fire obeys a Hypergeometric probability distribution.

To further permit the analysis, we shall divide the A force into two parts, that initially not under fire and that initially under fire. If we define the part of the A force under fire as

$$a_0 = \frac{A_0\, d}{L}, \qquad \text{(V.E-15)}$$

we may write the initial force as

$$A_0 = A_0 - B_0\, a_0 + B_0\, a_0. \qquad \text{(V.E-16)}$$

After n fires, but before any of the B force move their aim points, A_n has the form

$$A_n = A_0 - B_0\, a_0 + B_0\, a_n, \qquad \text{(V.E-17)}$$

where a_n is the part of a_0 remaining alive in the kill area of a fire. Let us further denote by f_n the fraction killed by the n^{th} fire. Thus

$$a_n = a_0 \prod_{i=1}^{n} (1 - f_i). \qquad \text{(V.E-18)}$$

Equation (V.D-13) gives the expected number of live targets selected per fire if we rewrite it as

$$f_n\, a_{n-1} = \frac{a_{n-1}\, f_1\, a_0}{a_0}, \qquad \text{(V.E-19)}$$

where:

a_{n-1} = number of living targets out of a_0,
a_0 = total number of targets in kill area,
$f_1\, a_0$ = number of targets killed.

This reduces to

$$f_n = f_1. \qquad \text{(V.E-20)}$$

Thus there is no difference in the fractional kill when we do not allow relocation from when we do. (Actually, if we applied this model to the case with relocation, the fractional kill would change. This is left as an exercise. What is significant here is that the same form of the kill rate can be found regardless of the limiting assumptions on the form of the engagement. The area fire is the critical and dominant factor.)

One of the questions that arises is whether the use of the expected value of kill for each fire is valid. To demonstrate that the assumption of independent firing is adequate, we calculate the expected loss from two repeated fires. While each fire is described by a Hypergeometric probability distribution, we may simplify the mathematics by requiring that the effective area of the fire be small (equivalently, that the expected loss per fire is small.) In this case, we may replace the Hypergeometric

probability distribution with a binomial probability distribution. The probability of n kills out of a population of N in the binomial probability distribution is

$$P(N : n) = \binom{N}{n} p^n q^{N-n}, \tag{V.E-21}$$

where: p is the probability of a kill, and q ≡ 1 - p. The loss from one fire is

$$<n> = \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} n, \tag{V.E-22}$$

however, it is easier to calculate the nonkilled fraction as

$$\begin{aligned} <N-n> &= \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} (N - n), \\ &\equiv A_1. \end{aligned} \tag{V.E-23}$$

It is relatively straightforward to find that this summation is

$$A_1 = q N = (1 - p) A_0. \tag{V.E-24}$$

If we now calculate the nonkilled after two fires, we may write

$$A_2 = \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} \sum_{n'=0}^{N-n} \binom{N-n}{n'} p^{n'} q^{N-n-n'} (N - n - n'), \tag{V.E-25}$$

where the second summation is the expected value for the second fire which is weighted by the probable nonkilled (not the expected value) of the first fire. We may rewrite this as

$$\begin{aligned} A_2 &= \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} q \frac{d}{dq} \sum_{n'=0}^{N-n} \binom{N-n}{n'} p^{n'} q^{N-n-n'}, \\ &= \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} q \frac{d}{dq} (p + q)^{N-n} \\ &= q \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} (N - n) \\ &= q^2 N = (1 - p)^2 N \\ &= (1 - p)^2 A_0, \end{aligned} \tag{V.E-26}$$

which confirms our contention.

V.F. Statistical Quadratic Lanchester Equation

The statistical quadratic case does not differ greatly from the deterministic quadratic case for the analyses that we shall present here. In our first analysis, we shall consider the B force to fire on a unit of the A force every Δt with a probability of kill p. We shall take as doctrine that a firer will continue to fire at a target until it is killed. If p is small, so that the likelihood of kill is small, the early (before many targets are killed) form of A , but after n shots, is

$$A_n = A_0 - B \sum_{i=0}^{n-1} (1 - p)^i p. \tag{V.F-1}$$

We may rewrite this as

$$A_n = A_0 - B p \sum_{i=0}^{n-1} e^{i \ln(1 - p)}, \tag{V.F-2}$$

and perform the summation

$$A_n = A_0 - B p \frac{1 - e^{n \ln(1 - p)}}{1 - e^{\ln(1 - p)}}, \tag{V.F-3}$$

which reduces to

$$A_n = A_0 - B \left[1 - e^{n \ln(1 - p)} \right]. \tag{V.F-4}$$

Since p is small, we may use Equation (V.E-10) to approximate this as

$$A_n = A_0 - B \left[1 - e^{-n p} \right]. \tag{V.F-5}$$

and change from integer to temporal form as

$$A_n = A_0 - B \left[1 - e^{-\frac{t p}{\Delta t}} \right]. \tag{V.F-6}$$

We may now take the derivative of A(t), with respect to time, and get

$$\frac{dA}{dt} = -\frac{p}{\Delta t} B e^{-\frac{t p}{\Delta t}}, \tag{V.F-7}$$

which has the form of the Quadratic Lanchester Differential Equation when t = 0, and we may identify the attrition rate as

$$\alpha = \frac{p}{\Delta t}, \qquad \text{(V.F-8)}$$

the result for the deterministic case. Unfortunately, this analysis leads us to the conclusion that the deterministic form of the Quadratic Lanchester Differential Equation is wrong. To resolve this, we must reexamine the analysis in a different light.

To resolve this incorrect form, we examine the process as being described by a Poisson probability distribution. (This relaxes the engage to kill restriction.) The B force fires B shots, each with individual probability of kill p, per time interval Δt. In Poisson probability distribution terms then, the probability of kill increment may be viewed in either of two ways: there is a probability of one kill of B p per time Δt, or there is a probability of one kill of p per time Δt/B.

We may write the probability of n kills over time t as

$$P_n(t) = \frac{\left(\frac{B\,p\,t}{\Delta t}\right)^n e^{-\frac{B\,p\,t}{\Delta t}}}{n!}, \qquad \text{(V.F-9)}$$

and the expected value of kills over time is

$$\langle n\rangle(t) = \frac{B\,p\,t}{\Delta t}. \qquad \text{(V.F-10)}$$

The mathematical form of the force strength of A is

$$A(t) = A_0 - B\,\frac{p\,t}{\Delta t}, \qquad \text{(V.F-11)}$$

which has the derivative,

$$\frac{dA}{dt} = -\frac{p}{\Delta t}\,B, \qquad \text{(V.F-12)}$$

which is exactly Lanchester's Quadratic Differential Equation.

We could equally well have represented this with a binomial probability distribution. The probability of n kills in time Δt is

$$p(n) = \binom{B}{n} p^n q^{B-n}, \tag{V.F-13}$$

which has the expected value

$$<n> = B\, p. \tag{V.F-14}$$

Thus,

$$\begin{aligned} A_1 &= A_0 - B\, p, \\ A_2 &= A_1 - B\, p \\ &= A_0 - 2\, B\, p, \end{aligned} \tag{V.F-15}$$

(we have previously demonstrated the independence here.)

$$A_n = A_0 - n\, B\, p. \tag{V.F-16}$$

This may be used to write the finite difference as

$$\frac{\Delta A}{\Delta t} = -\frac{p}{\Delta t} B, \tag{V.F-17}$$

which goes over to the time derivative directly since A(t) is jump discontinuous.

V.G. Mixed Law

Because the mixed Lanchester law describes combat between two forces whose individual attrition differential equations are linear and quadratic (individually,) in form, no separate Ironman Analysis is necessary doe to the inherent nature of relaxing attrition of one of the two forces involved in the Ironman Analysis. The analyses presented for the linear and quadratic Lanchester laws thus also apply to the mixed Lanchester Law.

V.H. Summary/Conclusions

These analyses are not presented as any rigorous proof of the validity of the Lanchester differential equations. Rather, they are intended to advance the concept that the Lanchester differential equations are definitions of the attrition rates rather than the opposite. In a later chapter, we shall examine the idea of a transport theory of combat. In that chapter, we shall consider the Lanchester differential equations both as scattering terms and as ensemble averages.

Before closing out the chapter, it is useful to consider the general from of the Lanchester differential equations, Equations (III.A-1). For our restrictions of $B_0 \equiv 1$, and $\beta = 0$, these become a single differential equation,

$$\frac{dA}{dt} = -\alpha\, A^{2-n}, \tag{V.H-1}$$

which we may rewrite as

$$\frac{dA^{n-1}}{dt} = -\alpha(n - 1). \tag{V.H-2}$$

This equation has the appearance of being invalid for n = 1, the Linear Law case.

If we solve Equation (V.H-2), we get

$$A(\tau)^{n-1} - A_0^{n-1} = -\alpha\ \tau\ (n - 1). \tag{V.H-3}$$

For our assignment of τ to be the time to kill one enemy unit, this equation may be rewritten as

$$\alpha\ \tau\ (n - 1) = A_0^{n-1} - (A_0 - 1)^{n-1}. \tag{V.H-4}$$

Since $A_0 \gg 1$ by assumption, we may rewrite and expand this equation,

$$\begin{aligned} \alpha\ \tau\ (n - 1) &= A_0^{n-1} - A_)^{n-1}\left(1 - \frac{1}{A_0}\right)^{n-1} \\ &\simeq A_0^{n-1} - A_)^{n-1}\left(1 - \frac{n-1}{A_0}\right) \\ &\simeq (n - 1)\ A_0^{n-2} \\ \alpha\ \tau &\simeq A_0^{n-2}, \end{aligned} \tag{V.H-5}$$

which is valid even for n = 1. This relation is useful since it gives us a scaling equation in attrition rate constants across attrition order (n). For example, if we designate the attrition rate for attrition order n by α_n, and the attrition rate for the Quadratic Law (for which Equation (V.H-5) reduces to Bonder's Equation,) as α_2, then Equation (V.H-5) gives us

$$\alpha_n = \alpha_2\, A_0^{n-2}. \tag{V.H-6}$$

Because α_2 is commonly associated with direct fire, Equation (V.H-6) allows us to scale attrition rates for other attrition orders from the (presumably known, or at least, easily calculated,) Quadratic Law attrition rates. In fact, we may substitute Equation (V.H-6) back into Equations (III.A-1) to get differential equations of the form,

$$\begin{aligned} \frac{dA}{dt} &= -\alpha_n A^{2-n} B \\ &= -\alpha_2 A_0^{n-2} A^{2-n} B \\ &= -\alpha_2 \left(\frac{A}{A_0}\right)^{2-n} B, \end{aligned} \tag{V.H-7}$$

which opens up a whole new arena of approximations. This will assume importance a we address the Osipov problem and subsequently.

1. Helmbold, Robert L., "Volley Fire Models", **Proceedings of the Workshop on Modeling and Simulation of Land Combat**, Leslie G. Callahan, Jr., ed., Calloway Gardens, GA, sponsored by the Georgia Institute of Technology, Atlanta, GA, 28-31 March 1982, pp. 287-301.

VI. MATHEMATICAL THEORY II: Further Solutions of the Lanchester Attrition Differential Equations

VI.A. Introduction

This chapter deals with a 'pot pourri' of mathematical topics associated with the Lanchester attrition differential equations in their simplest forms as initially solved in Chapter III. In some instances, these topics deal with the limiting cases of previously considered problems while other topics deal with approximations, or alternate views or approaches to the attrition differential equations which will have relevance as we consider more elaborate forms of the attrition differential equations than the pure forms hereto considered. This chapter thus serves as a mathematical introduction to topics which we shall encounter in later chapters. As before, the student who is not mathematically inclined has the option of accepting the derivations at face value, and need merely note the results for future use.

The first two topics in this chapter are the near-draw and draw solutions of the attrition differential equations. We recall the conclusion condition

$$\Delta_n = \alpha B_0^n - \beta A_0^n, \tag{VI.A-1}$$

which in mathematical terms describes the state of the "winner" if combat is carried to a conclusion. (If $\Delta_n > 0$, then the Blue force is the "winner" while if $\Delta_n < 0$, then the Red force is the winner.) If $\Delta_n = 0$ exactly, and combat is carried to a conclusion, then there is no "winner"; both sides are reduced to zero force strength. This is referred to as a draw condition or situation, and the attrition differential equations have special solutions.

The draw case solutions have, of course, already been sketched in Chapter IV, but they were derived there by expanding the general exact time solutions. In this chapter, the draw case solutions are derived directly from the attrition differential equations without recourse to expansions and limits as $\Delta_n \to 0$.

If Δ_n is small, then a condition of near-draw occurs and there are special (but not necessarily unique) approximate solutions to the attrition differential equations.

Another analytical (in the mathematical sense,) topic which we shall pursue is that of inverse solutions. Normally, we solve the attrition differential equations for explicit solutions of the force strengths as functions of time with initial force strengths and the attrition rates as parameters. Inverse solutions are explicit solutions of the attrition rates as functions of time with initial and final (or at least intermediary) force strengths as parameters. These solutions are useful in the analysis of historical data.

Next, we examine expressions of the attrition differential equations as attrition integral equations. This examination lays the basis for introducing approximate numerical techniques for the attrition differential equations

Finally, we deal with a combined attrition differential equation and the quadratic law differential equations with reinforcements as introductions to a class of attrition differential equations which either do not possess state solutions or if they do, the state solutions are so complex as to limit their use in arriving at exact solutions of the differential equations. An example of such are differential equations which possess transcendental state solutions. Such differential equations cannot be solved using the powerful method of normal forms described in Chapter III.

VI.B. Near-Draw Solutions

The near-draw situation arises when the conclusion condition is small. In mathematical terms,

$$
\begin{aligned}
\alpha B^n &> \Delta, \\
\beta A^n &> -\Delta,
\end{aligned}
\qquad \text{(VI.B-1)}
$$

which states that combat ends much before conclusion and normally refers to situations where the two forces are approximately evenly matched. This approximation may also be used to describe the early stages of combat before losses are too great. The general use of these solutions is to gain insight from differential equations which do not lend themselves to exact solutions. Note that the near-draw problem does not really have anything magic to do with &n being small; rather, it has to do with a short duration conflict in a mathematical sense.

If the near-draw situation holds, then we may usually take the normal form expression of the state solution,

$$
B = \left[\frac{\beta}{\alpha} A^n + \frac{\Delta_n}{\alpha} \right]^{\frac{1}{n}}, \qquad \text{(VI.B-2)}
$$

(or the equivalent expression in A as a function of B,) and expand it as follows:

$$
\begin{aligned}
B &= \left(\frac{\beta}{\alpha}\right)^{\frac{1}{n}} A \left[1 + \frac{\Delta_n}{\beta A^n} \right]^{\frac{1}{n}} \\
&\approx \left(\frac{\beta}{\alpha}\right)^{\frac{1}{n}} A \left[1 + \frac{\Delta_n}{n\ \beta\ A^n} \right] \\
&\approx \left(\frac{\beta}{\alpha}\right)^{\frac{1}{n}} \left[A + \frac{\Delta_n}{n\ \beta\ A^{n-1}} \right],
\end{aligned}
\qquad \text{(VI.B-3)}
$$

which effectively linearizes the normal form expression. We note immediately that for the linear law attrition case, n = 1, this expression does not exist since the linear law normal form expression is already linear. Thus, the linear law attrition differential equations do not possess near-draw solutions of this form.

VI.B.1 Near-Draw Linear Law Solutions

A near-draw approximate solution for the linear Lanchester law can be derived by expanding the exact solutions. The process is essentially that used before in Section IV.B.3 to derive the draw case solutions. In this instance, however, we use

the expansion,

$$e^{\pm\Delta_1 t} \simeq 1 \pm \Delta_1 \, t + \frac{\Delta_1^2 \, t^2}{2} \pm \ldots., \tag{VI.B-4}$$

and substitute it into the exact solutions, Equations (III.C.2) and (IV.A.14),

$$\begin{aligned} A(t) &= \frac{A_0 \, \Delta_1}{\beta A_0 - \alpha B_0 \, e^{-\Delta_1 t}} \\ B(t) &= \frac{B_0 \, \Delta_1}{\beta A_0 \, e^{\Delta_1 t} - \alpha B_0}, \end{aligned} \tag{VI.B-5}$$

to yield

$$\begin{aligned} A(t) &= \frac{A_0 \, \Delta_1}{\beta A_0 - \alpha B_0 \left(1 - \Delta_1 t + \dfrac{\Delta_1^2 t^2}{2}\right)} \\ B(t) &= \frac{B_0 \, \Delta_1}{\beta A_0 \left(1 + \Delta_1 t + \dfrac{\Delta_1^2 t^2}{2}\right) - \alpha B_0}, \end{aligned} \tag{VI.B-6}$$

These two equations may be rewritten as

$$\begin{aligned} A(t) &= \frac{A_0 \, \Delta_1}{\beta A_0 - \alpha B_0 + \alpha B_0 \left(\Delta_1 t - \dfrac{\Delta_1^2 t^2}{2}\right)} \\ B(t) &= \frac{B_0 \, \Delta_1}{\beta A_0 - \alpha B_0 + \beta A_0 \left(\Delta_1 t + \dfrac{\Delta_1^2 t^2}{2}\right)}, \end{aligned} \tag{VI.B-7}$$

We now use the definition of $\Delta_1 \equiv ß \, A_0 - \alpha \, B_0$, to further rewrite these equations as

$$A(t) = \frac{A_0\ \Delta_1}{\Delta_1 \ + \alpha B_0\left(\Delta_1 t - \dfrac{\Delta_1^2 t^2}{2}\right)}$$

$$B(t) = \frac{B_0\ \Delta_1}{\Delta_1 + \beta A_0\left(\Delta_1 t + \dfrac{\Delta_1^2 t^2}{2}\right)}, \qquad \text{(VI.B-8)}$$

and cancel terms Δ_1 in both numerator and denominator to get

$$A(t) = \frac{A_0}{1 + \alpha B_0\left(t - \dfrac{\Delta_1 t^2}{2}\right)}$$

$$B(t) = \frac{B_0}{1 + \beta A_0\left(t + \dfrac{\Delta_1 t^2}{2}\right)}, \qquad \text{(VI.B-9)}$$

Note that if $\Delta_1 > 0$ (Red "winner" at conclusion,) then the decrease in Red force strength, Equation (VI.B-9a) is retarded since the denominator increases in a slower fashion (due to the 'minus' Δ_1 term, while the decrease in Blue force strength, Equation (VI.b-9b) is accelerated since the denominator increases faster (due to the 'plus' Δ_1 term.) The exact opposite occurs when $\Delta_1 < 0$.

The behavior of this approximation is demonstrated in the series of Figures (VI.B-1) - (VI.B-5). Figures (VI.B-1) and (VI.B-2) present both the exact and the approximate near-draw time solutions for values of Δ_1 of ± 0.05, respectively. For this small value of Δ_1 little difference is evident in the two solutions. Figures (VI.B-3) and (VI.B-4) present the same curves, but for Δ_1 values of ± 0.5, an increase of one order of magnitude. Some error can be discerned in the curves at long time (relatively). If, however, we double the value of Δ_1 to -1, as shown in Figure (VI.A.1-5), we discern the instability of the approximation. This instability occurs in the Blue force strength since Δ_1 is negative. The cause of this instability can be seen by noting that the denominator of Equation (VI.B-9a) has a positive part equal to

$$1 + \beta A_0 t, \qquad \text{(VI.B-10)}$$

and a negative part equal to

$$\frac{\beta A_0 \Delta_1 t^2}{2}. \qquad \text{(VI.B-11)}$$

Since the magnitude of Δ_1 is one, the negative part may be written as

$$-\frac{\beta A_0 t^2}{2}. \qquad \text{(VI.B-12)}$$

As long as t is sufficiently small that the negative part is less than the positive part, the denominator is well behaved. Once the value of t gets bigger than one, however, the denominator actually begins to get smaller, and the calculated force strength becomes larger instead of smaller. This effect is clearly incorrect. From this behavior, we can calculate the limits on time of the behavior of the approximate solutions, Equations (VI.B-9). One of the denominators of these equations will have a maximum value when

$$t = |\Delta_1|^{-1}. \qquad \text{(VI.B-13)}$$

This value of time is the point where the approximation has definitely become incorrect, so use of the approximation should be limited to values of time which are a fraction of this value (say 1/3 or 1/2.)

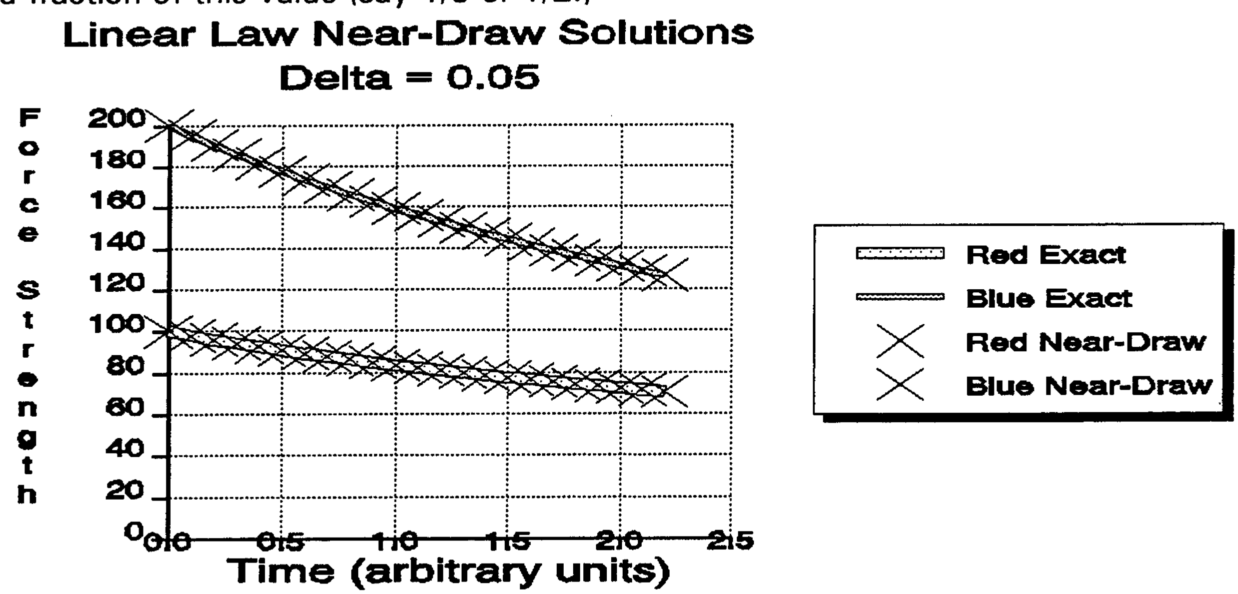


Figure VI.B-1

The student may note that this limitation on the validity of the approximation really is another statement of the smallness of Δ_1. The approximation is good only

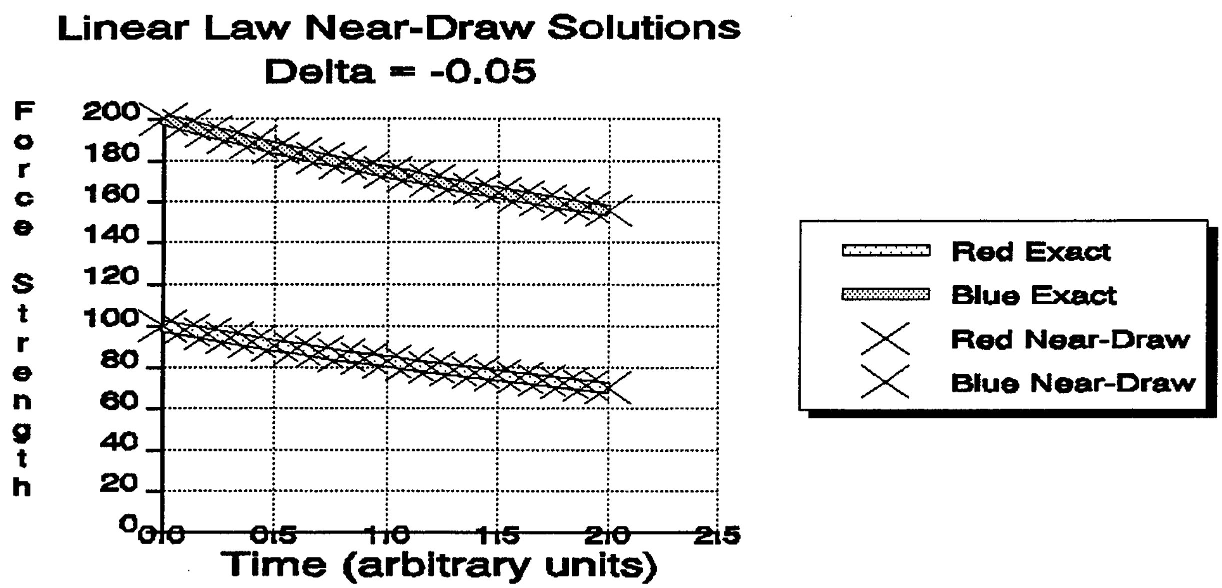


Figure VI.B-2

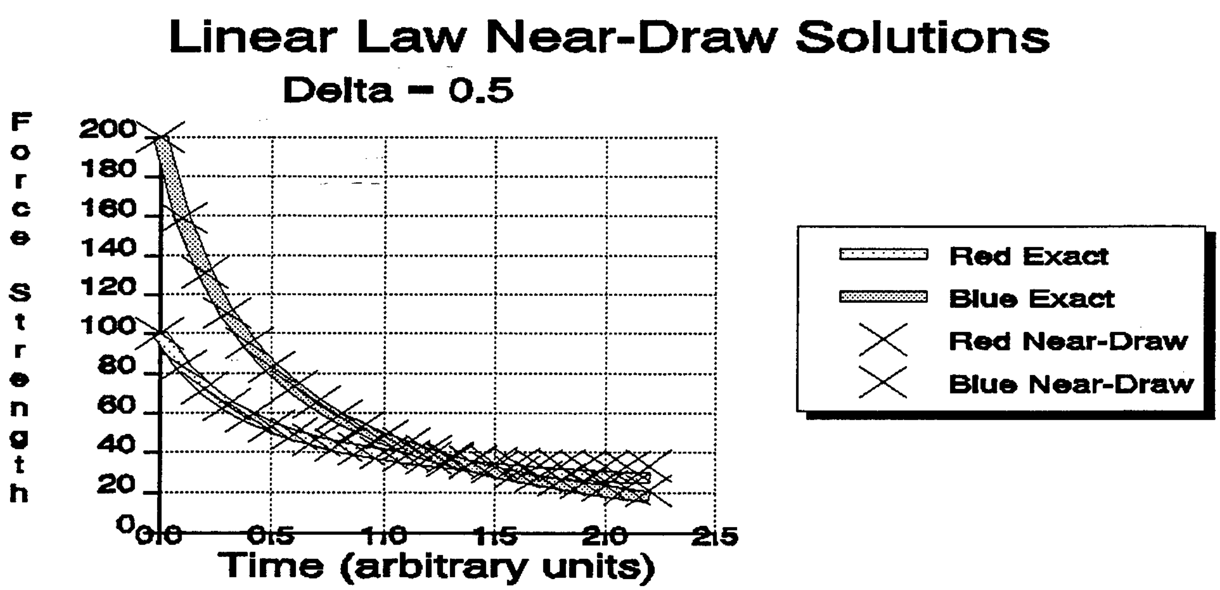


Figure VI.B-3

for small magnitude values of Δ_1. Since the expansion that was used to derive the approximation was based on the product of Δ_1 and t being small, it is readily seen that this product must remain small for the approximation to be valid. Thus there is a balance between the value of Δ_1 and of t. As long as the battle is short enough in duration, a large value for Δ_1 is admissible. Conversely, for a small value of Δ_1, long duration battles may be approximated.

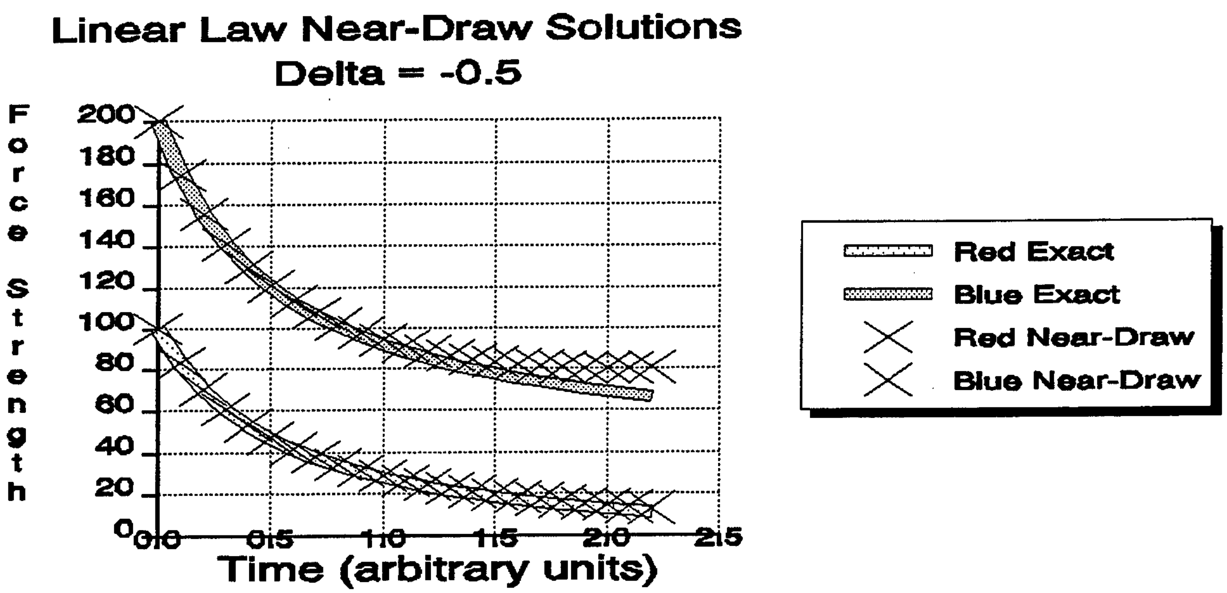


Figure VI.B-4

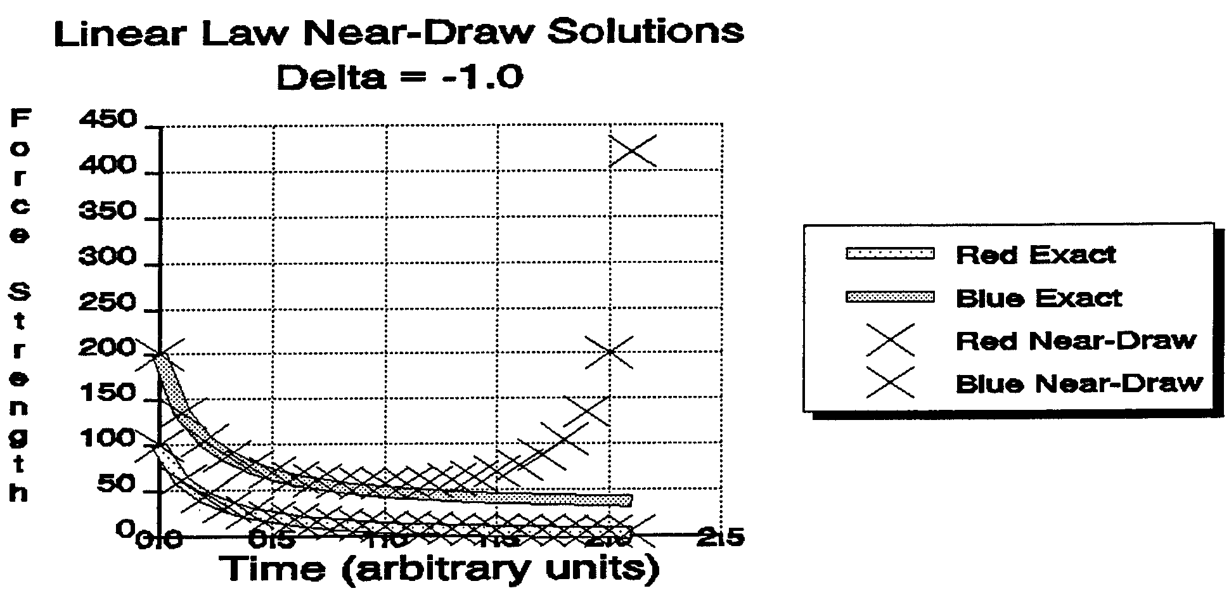


Figure VI.B-5

VI.B.2 Near-Draw Square Law Approximation

For the square law, Equation (VI.B-3) may be explicitly rewritten as

$$B = \sqrt{\frac{\beta}{\alpha}}\left(A + \frac{\Delta_2}{2\beta A}\right), \qquad \text{(VI.B-14)}$$

which we may more usefully rewrite as

$$B = \frac{1}{\delta}\,\frac{2\beta A^2 + \Delta_2}{2\beta A}, \qquad \text{(VI.B-15)}$$

where δ has been defined earlier in Chapter III.

If we now substitute Equation (VI.B-15) into the appropriate square law attrition differential equation, Equation (III.A-4),

$$\frac{dA}{dt} = -\alpha\, B,\ (III.A\text{-}3), \qquad \text{(VI.B-16)}$$

and rearrange the result slightly, we may write the exact differential equation

$$\frac{2\beta A\, dA}{2\beta A^2 + \Delta_2} = -\gamma\, dt. \qquad \text{(VI.B-17)}$$

This equation may be directly integrated since the numerator is the exact derivative (to within a constant) of the denominator,

$$d(2\beta A^2 + \Delta_2) = 4\beta A\, dA, \qquad \text{(VI.B-18)}$$

to yield

$$\ln(2\beta A^2 + \Delta_2)\Big|_{A_0}^{A(t)} = -2\gamma t. \qquad \text{(VI.B-19)}$$

The limits may be applied, and the antilogarithm taken,

$$A(t)^2 + \frac{\Delta_2}{2\beta} = A_0^2 + \frac{\Delta_2}{2\beta}\, e^{-2\gamma t}. \qquad \text{(VI.B-20)}$$

We may rewrite Equation (VI.A.2.3) as

$$A(t)^2 = A_0^2\, e^{-2\gamma t} - \frac{\Delta_2}{2\beta}\left[\,1 - e^{-2\gamma t}\,\right]. \qquad \text{(VI.B-21)}$$

The functional form for A(t) may be formed by expanding the square root since Δ_2 is, by assumption, small (the near-draw situation) to yield

$$A(t) = A_0\, e^{-\gamma t} - \frac{\Delta_2}{2\beta A_0^2}\, \sinh(\gamma t). \qquad \text{(VI.B-22)}$$

(Alternatively, if t is small, the first left hand side terms is larger than the second, regardless of the value of Δ_2.) This equation, (VI.B-22) is the near draw approximation of the square law for the Red force strength. The equivalent equation for the Blue force strength may immediately be written using symmetry as

$$B(t) = B_0\, e^{-\gamma t} + \frac{\Delta_2}{2\alpha B_0^2}\, \sinh(\gamma t). \qquad \text{(VI.B-23)}$$

We again compare the exact and approximate solutions in a series of figures, numbers (VI.B-6) - (VI.B-8). In the first figure, values of $\Delta_2 = -200$, and $\gamma = 0.14$ are used. Quite good agreement can be seen despite the large magnitude of Δ_2. (Remember, it is the size of Δ_2 relative to the final force power that is important here, since we are allowing t to become large.) Somewhat worse agreement at long time may be seen in Figure (VI.B-7) where values of $\Delta_2 = -2000$ and $\gamma = 0.14$ are used. We see however, that catastrophe occurs in Figure (VI.B-8). The catastrophe occurs when the second terms in Equations (VI.B-22) and (VI.B-24) become larger than the first terms. Mathematically, the catastrophe occurs since $\sinh(\gamma t)$ contains a term $\propto e^{\gamma t}$. The catastrophe may be avoided by ensuring that the approximations are not used when the second terms are large compared to the first terms.

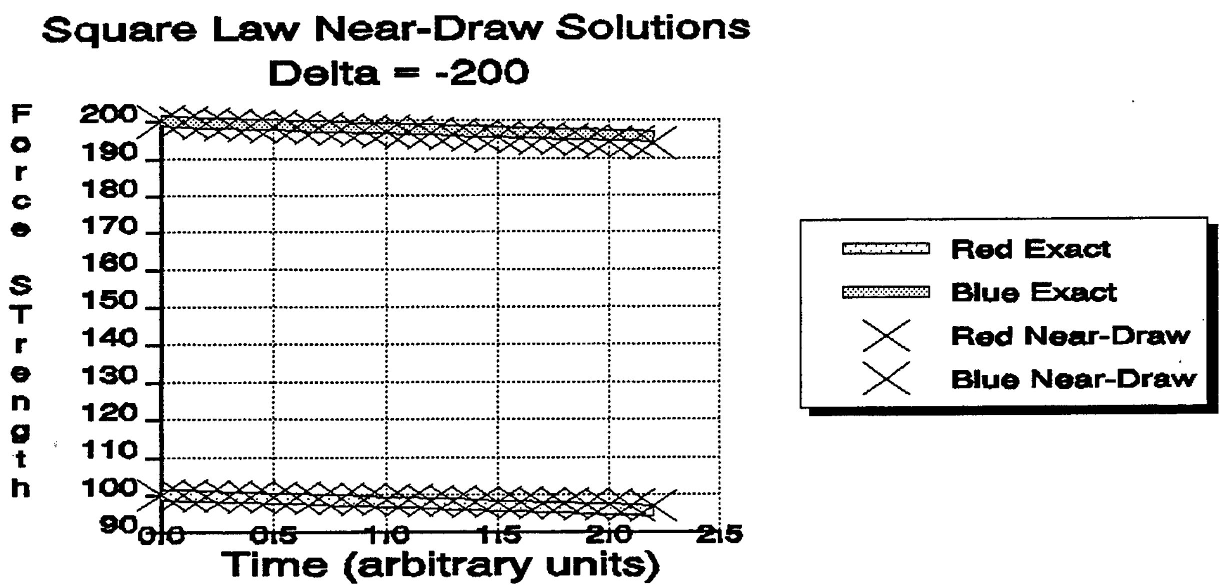


Figure VI.B-6

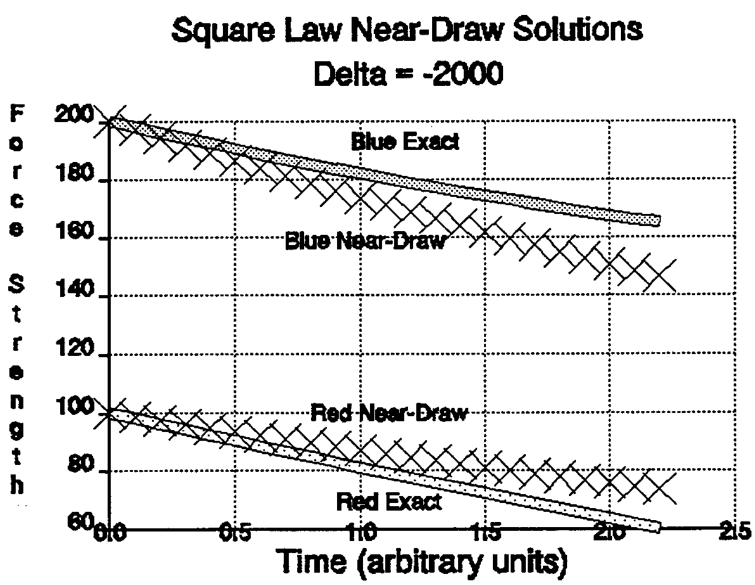


Figure VI.B-7

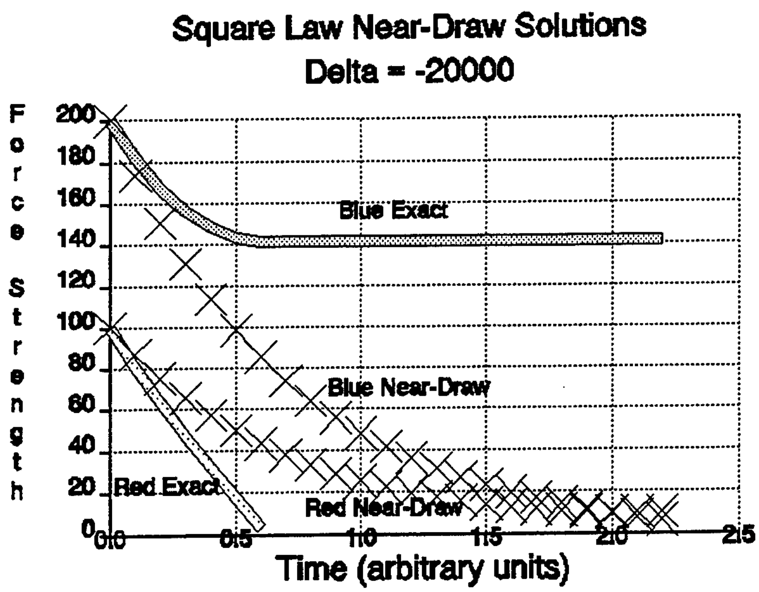

Square Law Near-Draw Solutions
Delta = -20000
Force Strength
200
180
160
140
120
100
80
60
40
20
0
Blue Exact
Blue Near-Draw
Red Exact
Red Near-Draw
0.0
0.5
1.0
1.5
2.0
2.5
Time (arbitrary units)

Figure VI.B-8

VI.B.3 Near-Draw Mixed Law Approximate Solutions

As we would expect from what we have seen in the two preceding sections on linear and square law near-draw approximations, expansion of the conclusion condition for the mixed law is a mixed approach. The conclusion condition cannot be expanded for the Blue force since the conclusion condition is already linear in Blue force strength. Substitution of the conclusion condition in the Red force attrition differential equation thus would result in the exact differential equation whose solution was developed in Chapter III. This is exactly what we would expect from the Red force attrition differential equation being linear-like.

Alternately, the conclusion condition for the Red force strength can be expanded and substituted into the Blue force attrition differential equation as an approximation. This is not a desirable prescription however, as it would result in an exact Red force strength solution and an approximate Blue force strength solution. A more useful approach is to proceed in the same manner that we sketched the mixed law draw case approximation in Section IV.D. In this case we use the expansion

$$\sec(x) \approx 1 + \frac{x^2}{2}, \qquad \text{(VI.B-24)}$$

and proceed as before in Section IV.D. This allows us to write

$$A(t) \approx \frac{A_0 - \Delta_m t}{1 + \beta A_0 t}, \qquad \text{(VI.B-25)}$$

and

$$B(t) \approx \frac{B_0 \left(1 + \dfrac{\beta \Delta_m t^2}{2} \right)}{(1 + \beta A_0 t)^2}. \qquad \text{(VI.B-26)}$$

We do not present figures depicting the behavior of these solutions, leaving them as an exercise for the student, but do note that care must be taken that the numerators of Equations (VI.B-25) and (VI.B-26) remain strictly positive (which depends on the sign of Δ_m.)

VI.C Draw Solutions

The draw situation arises when the conclusion condition is exactly zero. In other words, when the fighting power of the two forces are exactly equal. Unlike the near-draw solutions presented in the previous section, which are approximations and therefore not unique, the draw solutions are unique and exact, and as we shall examine later in the chapter on historical insight, enjoy a special place in Lanchester theory.

As we have noted before, the draw situation is also a special case from an attrition, as well as a mathematical standpoint. If combat for the draw situation is carried to a conclusion, neither side will have any units remaining. This does not necessarily mean that the forces of both sides would be zero, but rather that they have no units capable of fighting. This is, however, a philosophical concept that we will not pursue further here. It is sufficient to assume that the mathematical interpretation is valid - that there are no units left on either side.

(I am always reminded whenever I consider the draw solutions of a cartoon drawn by Gahan Wilson some years ago which shows a sole surviving soldier in CBR gear amidst the ruins of a battlefield, shouting "We won - I think.")

VI.C.1 Quadratic Law Draw Solution

(We depart here from the usual order: linear-quadratic-mixed; to allow the natural introduction of an approximation to the draw solutions for the linear and mixed law cases.)

In the draw case for the quadratic law, the state solution reduces to

$$\alpha B^2 = \beta A^2, \tag{VI.C-1}$$

which we may simplify as

$$B = \frac{A}{\delta}, \tag{VI.C-2}$$

and substitute into Equation (III.A-4)

$$\frac{dA}{dt} = -\alpha\ B, \tag{VI.C-3}$$

(and its conjugate,) to get

$$\frac{dA}{dt} = -\gamma A. \tag{VI.C-4}$$

This differential equation has the direct (and immediate) solution

$$A(t) = A_0 e^{-\gamma t}, \tag{VI.C-5}$$

with a conjugate solution

$$B(t) = B_0 e^{-\gamma t}. \tag{VI.C-6}$$

We may note immediately that in the draw situation, quadratic law combat is infinite in duration. If we compare this attrition process with any other quadratic law attrition, we find the draw situation to take the longest time. This factor will be of considerable impact when we investigate the historical insights of attrition.

VI.C.2 Linear Law Draw Solution

For the linear law draw case, the state solution reduces to

$$B = \frac{\beta}{\alpha} A, \tag{VI.C-7}$$

which may be substituted into the appropriate attrition differential equation, equation (III.A-2), to yield

$$\frac{dA}{dt} = -\beta A^2. \tag{VI.C-8}$$

This differential equation has the solution

$$A(t) = \frac{A_0}{1 + \beta A_0 t}. \tag{VI.C-9}$$

The solution for the Blue force strength may be formed using either symmetry, or from the same prescription as above, as

$$B(t) = \frac{B_0}{1 + \alpha B_0 t}. \tag{VI.C-10}$$

Note that because of the form of equation (VI.C-1), the denominators of equations

(VI.C-9) and (VI.C-10) are equal. Further, as was the case with the square law draw solution, the linear law draw situation battle is of infinite duration if fought to a conclusion. (Note that equations (VI.C-9) and (VI.C-10) follow directly from equations (VI.B-9) when $\Delta_1 = 0$.)

Earlier, we mentioned that the square law draw solutions would be developed first so as to naturally admit an approximate solution for the other two draw solutions. In the previous section, we derived the square law draw solutions which were simple exponentials in form. If we make use of the logarithm approximation introduced in the preceding chapter,

$$\ln(1 + x) \approx x,\ x\ small, \tag{VI.C-11}$$

then we may introduce an approximate form of equations (VI.C-9) and (VI.C-10), for small value of t,

$$A(t) \approx A_0\, e^{-\beta A_0 t}, \tag{VI.C-12}$$

and

$$B(t) \approx B_0\, e^{-\alpha B_0 t}. \tag{VI.C-13}$$

These two equations are known as the exponential approximations for the linear law draw solutions. They are useful in comparisons among the three draw solution sets.

VI.C.3 Mixed Law Draw Solutions

For the draw solution of the mixed law, the state solution reduces to

$$2\,\alpha\, B = \beta\, A^2, \tag{VI.C-14}$$

which may be substituted into the mixed law attrition differential equations, equations (III.A-6) and (III.A-7) to yield

$$\frac{dA}{dt} = -\frac{\beta}{2}\, A^2,\ (square\text{-}like) \tag{VI.C-15}$$

and

$$\frac{dB}{dt} = -\sqrt{2\alpha\beta}\, B^{\frac{3}{2}}, \tag{VI.C-16}$$

We note that the square-like attrition differential equation, equation (VI.C-15), has the same form as the linear law draw situation differential equation (VI.C-8), but differs

in the factor of ½. This differential equations has the solution

$$A(t) = \frac{A_0}{1 + \frac{\beta A_0 t}{2}}. \tag{VI.C-17}$$

Thus the rate of loss of Red units in this square-like solution is one-half of that in the linear law case.

The solution of the linear-like differential equation, equation (VI.C-15), is

$$B(t) = \frac{B_0}{\left(1 + \frac{\gamma\sqrt{B_0}t}{\sqrt{2}}\right)^2}, \tag{VI.C-18}$$

where γ has the same definition introduced in Chapter III.

These two solutions have the exponential approximations

$$A(t) \approx A_0\, e^{-\frac{\beta A_0 t}{2}}, \tag{VI.C-19}$$

which has one-half the rate of attrition of the linear law exponential approximation, and

$$B(t) \approx B_0\, e^{-\gamma\sqrt{2B_0}t} \tag{VI.C-20}$$

Note that the mixed law attritions proceed according to different time scales.

As in the square and linear law draw cases, battle to a conclusion for the mixed law draw situation has infinite duration.

VI.D Inverse Solution

Under normal conditions, the object of interest is the solution set of the Lanchester attrition differential equation with the initial force strengths as (boundary) conditions and the attrition rates as parameters. In this section, the object of interest is the square law solution set with initial and "final" force strengths as parameters to provide functional solutions for the attrition rates. We limit ourselves to the square law because the linear law is transcendental in the initial force strengths and attrition rates. Further, we also limit ourselves to constant attrition rates.

The square law time solutions may be written as

$$A(t) = A_0 \cosh(\gamma t) - \delta B_0 \sinh(\gamma t), \qquad \text{(VI.D-1)}$$

and

$$B(t) = B_0 \cosh(\gamma t) - \frac{A_0}{\delta} \sinh(\gamma t), \qquad \text{(VI.D-2)}$$

where: $\gamma \equiv \sqrt{(\alpha ß)}$, and
$\delta \equiv \sqrt{(\alpha/ß)}$.

If we assume that the initial force strengths A_0 and B_0, and the force strengths $A(\tau)$ and $B(\tau)$ at some time τ are given, then these two equations have two unknown, γ and δ (or equivalently α and ß.)

By eliminating δ, we may find after a bit of algebra, that

$$\cosh(\gamma \tau) = \frac{A_0 B_0 + A(\tau)B(\tau)}{A_0 B(\tau) + A(\tau)B_0}. \qquad \text{(VI.D-3)}$$

The student will note that this quantity is always positive, but may be infinite when $A(\tau)$ and $B(\tau)$ are zero; at the conclusion of a draw battle. If we next exploit the definition of the arc cosh function,

$$\cosh^{-1}(x) = \ln\left(x + \sqrt{x^2 - 1}\right), \qquad \text{(VI.D-4)}$$

the above equation may be further reduced to

$$\gamma \tau = \ln\left(A_0 B_0 + A(\tau)B(\tau) + \sqrt{\left(A_0^2 - A(\tau)^2\right)\left(B_0^2 - B(\tau)^2\right)}\right), \qquad \text{(VI.D-5)}$$

which allows us to solve for γ.

Similarly, some algebra also allows us to write

$$\delta^2 = \frac{A_0^2 - A(\tau)^2}{B_0^2 - B(\tau)^2}, \qquad \text{(VI.D-6)}$$

which is just a restatement of the state solution.

Equations (VI.D-5) and (VI.D-6), with the definitions of γ and δ, may be used to calculate values of α and ß given the initial and final (or even intermediate) force strengths and the duration of the conflict. We shall return to these equations in the chapter on historical insights of attrition.*

---

* The student with access to the limited DoD literature may wish to compare these equations with those derived by Robert L. Helmbold in "Lanchester Parameters for Some Battles of the last two hundred years", Combat Operations Research Staff Paper CORG-SP-122, 14 February 1961, AD481201, LIMITED. Because this book is limited to the "open" literature, we cannot include information that is either classified or limited, as Dr. Helmbold's report is.

VI.E. Integral Equation Formalism

So far, we have considered only analytical solutions of the attrition differential equations and attrition rates which are constants. In subsequent chapters of this book, we shall encounter attrition differential equations for which we cannot find exact, analytical solutions and attrition rates which are not constant. Further, all of the Lanchester combat models that we have considered thus far have been for homogeneous forces (where the two opposing forces could be considered mathematically as only two homogeneous collections of units.) In subsequent chapters, we shall also consider heterogeneous forces.

In this section, we introduce an alternate formalism for the attrition process where we replace the differential equations with integral equations. This representation, while exactly equivalent to the differential equation representation, serves several purposes: it facilitates the consideration of attrition processes for which we cannot necessarily find exact analytical solutions such as the many problems with heterogeneous forces and/or variable attrition rates, and it provides a natural basis for the discussion and development of numerical calculation methods which may be used with digital computers to find approximate solutions. (These are introduced in the next section.)

To introduce the integral equation representation, we first write the quadratic law attrition differential equation in the form

$$\frac{d}{dt}\begin{pmatrix} A \\ B \end{pmatrix} = -\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix}, \tag{VI.E-1}$$

where:

$$\begin{pmatrix} A \\ B \end{pmatrix} \tag{VI.E-2}$$

is an array of force strengths and

$$\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \tag{VI.E-3}$$

is an array of attrition rate constants/functions.
As a shorthand, we will denote the force strength array as [F], and the attrition rate array as [ζ]. This allows the square law attrition differential equation to be written in matrix form as

$$\frac{d}{dt}[F] = -[\zeta][F]. \qquad \text{(VI.E-4)}$$

(Whenever possible matrix notation such as this will be used to keep the notation compact. Only when some point of the exposition necessitates it will we explicit write the matrix equations in element form.)

We next directly integrate both sides of equation (VI.E-1) with respect to time, and rearrange the result slightly (taking advantage of the fact that the left hand side of equation (VI.E-1) is an exact differential.) This gives

$$\begin{pmatrix} A(t+\Delta t) \\ B(t+\Delta t) \end{pmatrix} = \begin{pmatrix} A(t) \\ B(t) \end{pmatrix} - \int_t^{t+\Delta t} \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \begin{pmatrix} A(t') \\ B(t') \end{pmatrix} dt'. \qquad \text{(VI.E-5)}$$

This equation is the (exact) quadratic law attrition integral equation. In array (or matrix) notation, this may be written as

$$[F(t+\Delta t)] = [F(t)] - \int_t^{t+\Delta t} [\zeta]\,[F(t')]\,dt'. \qquad \text{(VI.E-6)}$$

If we let $t + \Delta t = t'$, since Δt only indicates some increment of time, equation (VI.E-5) may be substituted into itself, yielding (in the matrix notation of equation (VI.E-6), which we shall hereafter use interchangeably.)

$$\begin{aligned} [F(t+\Delta t)] = [F(t)] &- \int_t^{t+\Delta t} [\zeta]\,[F(t)]\,dt' \\ &+ \int_t^{t+\Delta t} [\zeta]\,dt' \int_t^{t'} [\zeta]\,[F(t'')]\,dt''. \end{aligned} \qquad \text{(VI.E-7)}$$

The first integral on the right hand side contains a force strength array which is independent of the variable of integration, and indeed, is equal to the force strength array which is the leading term on the right hand side of the equation. As a result, we may rewrite this equation as

$$\begin{aligned} [F(t+\Delta t)] = &\left([1] - \int_t^{t+\Delta t} [\zeta]\,dt' \right) [F(t)] \\ &+ \int_t^{t+\Delta t} [\zeta]\,dt' \int_t^{t'} [\zeta]\,[F(t'')]\,dt'', \end{aligned} \qquad \text{(VI.E-8)}$$

where [1] is the identity matrix.

At this point, we make use of a delicate point in the integration of matrices. Students who are unfamiliar with matrix algebra may wish to consult a text book on this subject. The delicate point is that we may bring the integral inside the matrix in

the first term since that integral has only one term. Thus

$$\int_t^{t+\Delta t} [\zeta]\, dt' = \int_t^{t+\Delta t} \begin{pmatrix} 0 & \alpha \\ \beta & p \end{pmatrix} dt'$$
$$= \begin{pmatrix} \int_t^{t+\Delta t} 0\, dt' & \int_t^{t+\Delta t} \alpha\, dt' \\ \int_t^{t+\Delta t} \beta\, dt' & \int_t^{t+\Delta t} 0\, dt' \end{pmatrix}, \tag{VI.E-9}$$

obviously, all the integrals over zeroes are themselves zero since the integrals are definite. If we denote the integrals over the attrition rates as

$$\underline{\alpha}(t + \Delta t) - \underline{\alpha}(t) = \int_t^{t+\Delta t} \alpha(t')\, dt', \tag{VI.E-10}$$

and

$$\underline{\beta}(t + \Delta t) - \underline{\beta}(t) = \int_t^{t+\Delta t} \beta(t')\, dt', \tag{VI.E-11}$$

This notation (the underscore) takes into account that the attrition rates may be functions of time. We also make use of the finite difference operator Δ to further write these attrition rate integrals as $\Delta\underline{\alpha}$(t) and $\Delta\underline{\beta}$(t), where the finite difference operator is defined by

$$\Delta f(t) \equiv f(t + \Delta t) - f(t). \tag{VI.E-12}$$

The integral over the attrition rate matrix may then be written as

$$\int_t^{t+\Delta t} [\zeta]\, dt' = \begin{pmatrix} 0 & \Delta\underline{\alpha}(t) \\ \Delta\underline{\beta}(t) & 0 \end{pmatrix}$$
$$\equiv \Delta[\underline{\zeta}](t), \tag{VI.E-13}$$

the integral equation may now be written

$$[F(t + \Delta t)] = \left([1] - \Delta[\underline{\zeta}](t) \right) [F(t)]$$
$$+ \int_t^{t+\Delta t} [\zeta](t')\, dt' \int_t^{t'} [\zeta](t'')\, [F(t'')]\, dt''. \tag{VI.E-14}$$

The substitution process may be repeated indefinitely to produce even more leading terms which are not integrals and a final term which is comprised of n

repeated integrals each over the attrition rate matrix times (finally) the force strength matrix, where n-1 is the number of substitutions of the primitive integral equation, equation (VI.E-5).

The linear law attrition differential equations may be converted to integral equations in a similar manner except that we start by writing the differential equations, equations (III.A-2) and (III.A-3) as

$$\frac{1}{A}\,\frac{dA}{dt} = -\alpha\, B, \qquad \text{(VI.E-15)}$$

and

$$\frac{1}{B}\,\frac{dB}{dt} = -\beta\, A, \qquad \text{(VI.E-16)}$$

These differential equations may be written in matrix form as

$$\begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} \frac{d}{dt} \begin{pmatrix} A \\ B \end{pmatrix} = - \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}, \qquad \text{(VI.E-17)}$$

or

$$[F^{-1}]\,\frac{d}{dt}\,[F] = -[\zeta]\,[F], \qquad \text{(VI.E-18)}$$

where $[F^{-1}]$ is the diagonal matrix of inverse force strengths. At this point, we must introduce the concept of the inverse of a (square) matrix, defined by the equation

$$[A]^{-1}\,[A] = [A]\,[A]^{-1} = [1], \qquad \text{(VI.E-19)}$$

where, as before, [1] is the identity matrix. The inverse matrix that we are interested in here is $[F^{-1}]^{-1}$ given by

$$[F^{-1}]^{-1} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \qquad \text{(VI.E-20)}$$

The student may confirm that this is indeed the inverse of $[F^{-1}]$.

We may now write equation (VI.E-18) as

$$\frac{d}{dt}[F] = -[F^{-1}]^{-1}[\zeta][F]. \tag{VI.E-21}$$

Using the same integration prescription as for the quadratic equation, this becomes

$$[F(t+\Delta t)] = [F(t)] - \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}[\zeta][F(t')]\,dt', \tag{VI.E-22}$$

which may in turn be cross substituted to yield

$$\begin{aligned}[F(t+\Delta t)] = [F(t)] &- \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}[\zeta][F(t)]\,dt' \\ &+ \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}[\zeta]\,dt' \int_t^{t'} [F^{-1}(t'')]^{-1}[\zeta][F(t'')]\,dt''.\end{aligned} \tag{VI.E-23}$$

We may again perform the first integral as before except that the integration must be performed on the individual elements of the product of the two arrays $[F^{-1}]^{-1}$ and $[\zeta]$. That is

$$\begin{aligned}\Delta\left([\underline{F^{-1}}]^{-1}[\underline{\zeta}]\right)(t) &= \int_t^{t+\Delta t} dt'\,[F^{-1}]^{-1}(t')\,[\zeta] \\ &= \begin{pmatrix} \int_t^{t+\Delta t} 0\,dt' & \int_t^{t+\Delta t} \alpha B(t')\,dt' \\ \int_t^{t+\Delta t} \beta A(t')\,dt' & \int_t^{t+\Delta t} 0\,dt' \end{pmatrix} \\ &\equiv \begin{pmatrix} 0 & \Delta\,\underline{\alpha B}(t) \\ \Delta\,\underline{\beta A}(t) & 0 \end{pmatrix},\end{aligned} \tag{VI.E-24}$$

where:

$$\begin{aligned}\Delta\,\underline{\alpha B}(t) &\equiv \int_t^{t+\Delta t} \alpha\; B(t')\,dt' \\ \Delta\,\underline{\beta A}(t) &\equiv \int_t^{t+\Delta t} \beta\; A(t')\,dt'.\end{aligned} \tag{VI.E-25}$$

The attrition rates may, again, be functions of time. Equation (VI.E-23) may now be rewritten as

$$\begin{aligned}[F(t+\Delta t)] = &\left([1] - \Delta\left([\underline{F^{-1}}]^{-1}[\underline{\zeta}]\right)(t)\right)[F(t)] \\ &+ \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}[\zeta]\,dt' \int_t^{t'} [F^{-1}(t'')]\,[\zeta]\,[F(t'')]\,dt''.\end{aligned} \tag{VI.E-26}$$

The mixed equations may be transformed in the same manner: the differential equations may be written in the form

$$\begin{pmatrix} 1 & 0 \\ 0 & B^{-1} \end{pmatrix} \frac{d}{dt} \begin{pmatrix} A \\ B \end{pmatrix} = - \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}, \qquad \text{(VI.E-27)}$$

or

$$[F_B^{-1}] \frac{d}{dt} [F] = -[\zeta] [F] \, . \qquad \text{(VI.E-28)}$$

The matrix integral equations for the mixed law thus looks like equation (VI.E-17), except that $[F^{-1}]^{-1}$ is replaced by $[F_B^{-1}]^{-1}$ (or $[F_A^{-1}]^{-1}$ if the Red force is linear-like.)

VI.F. Numerical Approximation

Generally, we do not know the analytic form of the force strengths although we will normally know the analytic forms of the attrition rates if they are functions (assumed of time.) Thus the introduction of the integral equations does not inherently offer any advantage in finding analytical solutions for the force strengths. Frequently, the attrition differential equations are sufficiently complex that they cannot be solved directly. This is especially true if heterogenous forces or attrition rate functions are involved. In these cases, approximate solutions must be found, usually in numerical form using a digital computer.*

The student should not be daunted by the need for a computer, however. Viable attrition simulations applying these numerical methods of solution may be formed from the attrition models in this book using spreadsheet programs such as LOTUS 1-2-3 (R). Most of the calculations in this book were performed in spreadsheet simulations using LOTUS. The use of such spreadsheet simulations is multiply handy since the entire simulation is readily open to inspection - the internal workings and equations are not hidden in arcane computer code; the input boundary conditions, attrition parameters, and the attrition equations themselves are available for easy modification and examination of excursions; and, finally, the calculated solutions are easily displayed graphically using the spreadsheet program's graphics capabilities. The student who is familiar with these programs but may not find the exposition here sufficiently clear may wish to refer to Orvis' book[1] which discusses the solution of differential equations (and other higher mathematical problems) using LOTUS.

In the preceding section, we derived the matrix integral equation for the quadratic law as

$$[F(t + \Delta t)] = [F(t)] - \int_t^{t+\Delta t} [\zeta]\, [F(t')]\, dt'. \qquad \text{(VI.F-1)}$$

The matrix integral equation for the linear law is

$$[F(t + \Delta t)] = [F(t)] - \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}\, [\zeta]\, [F(t')]\, dt'. \qquad \text{(VI.F-2)}$$

These equations readily lend themselves to numerical integration if we introduce one of a variety of integration approximations. Inherent to all of these approximations is

* Actually, an analog computer can be used, but analog computers are relatively rare today while digital computers are relatively common. See Leslie G. Callahan, Jr., and Glenn Crosby, "Lanchester Modeling of Small-Unit Combat, U.S. Army Missile Command Technical Report RD-CR-83-17, November 1981, AD8076783, LIMITED.

the requirement that some time interval Δt be set to a constant. The value of &t will be the critical factor in the accuracy of the approximate integration since all of the approximations we will use in this section will have errors of order Δt^2. More complicated (that is, higher order in Δt) integration approximations may be used in a bootstrap manner, but these will not be discussed here as they bring much complexity an little insight into the approximation process. The student who is interested in these higher order approximations is referred to a standard text on numerical approximation such as Carnahan, Luther, and Wilkes.[2]

It is not our purpose here to provide a textbook on numerical methods as applied to attrition mechanics, but rather to sketch the application of simple numerical methods which will be of use to the student in performing simple, but hopefully insightful calculations (such as spreadsheet calculations) in an independent manner.

The first integral approximation that we introduce is the zeroth order or rectangular rule. In this case, the function under the integral is approximated by its value at the lower (or upper) limit of the integral. Thus

$$\int_t^{t+\Delta t} f(t')\, dt' \approx f(t)\, \Delta t. \tag{VI.F-3}$$

This approximation is called the rectangular rule because the area represented by the integral is approximated by a rectangle of height f(t) and width Δt.

If we use this approximation, the matrix quadratic equation becomes

$$[F(t+\Delta t)] \approx ([1] - [\zeta(t)]\, \Delta t\,)\, [F(t)]. \tag{VI.F-4}$$

It is useful, for the homogeneous force case, to write the individual elements of this matrix equation as

$$A(t + \Delta t) \approx A(t) - \alpha(t)\, B(t)\, \Delta t, \tag{VI.F-5}$$

and

$$B(t + \Delta t) \approx B(t) - \beta(t)\, A(t)\, \Delta t, \tag{VI.F-6}$$

Whence here, and later, we have (and will usually,) assume the attrition rates to be functions.

Equations (VI.F-5) and (VI.F-6) may be directly substituted into simple simulations such as an electronic spreadsheet.

The matrix linear law integral equation may also be approximated using this

integration rule as

$$[F(t+\Delta t)] \approx ([1] - [F^{-1}(t)]^{-1} [\zeta(t)] \Delta t) [F(t)]. \qquad \text{(VI.F-7)}$$

The individual elements of this equation, for the homogeneous force case, may be written as

$$\begin{aligned} A(t + \Delta t) &\approx A(T) - A(t)\ \alpha(t)\ \Delta t\ B(t) \\ &= (1 - \alpha(t)\ B(t)\ \Delta t\ ) A(t), \end{aligned} \qquad \text{(VI.F-8)}$$

and

$$\begin{aligned} B(t + \Delta t) &\approx B(T) - B(t)\ \beta(t)\ \Delta t\ A(t) \\ &= (1 - \beta(t)\ A(t)\ \Delta t\ ) B(t). \end{aligned} \qquad \text{(VI.F-9)}$$

Note that both the quadratic and linear law force strength approximations, equations (VI.F-4)-(VI.F-9), can be solved directly.

A more complicated integral approximation is the first order or trapezoid rule. In this case, the function is approximated by a straight line (first order Taylor's series.) Thus

$$\int_t^{t+\Delta t} f(t')\ dt' \approx \frac{f(t+\Delta t) + f(t)}{2}. \qquad \text{(VI.F-10)}$$

If we use this rule, the matrix quadratic integral equation becomes

$$[F(t+\Delta t)] \approx [F(t)] - \frac{([\zeta(t+\Delta t)]\ [F(t+\Delta t)] + [\zeta(t)]\ [F(t)]) \Delta t}{2}. \qquad \text{(VI.F-11)}$$

which may be rewritten after some rearrangement as

$$\left([1] + \frac{[\zeta(t+\Delta t)]\ \Delta t}{2} \right) [F(t+\Delta t)] \approx \left([1] - \frac{[\zeta(t)]\ \Delta t}{2} \right) [F(t)]. \qquad \text{(VI.F-12)}$$

If we calculate the inverse of the leading term matrix on the left hand side of this equation, this is just

$$[F(t+\Delta t)] \approx \left([1] + \frac{[\zeta(t+\Delta t)]\ \Delta t}{2} \right)^{-1} \left([1] - \frac{[\zeta(t)]\ \Delta t}{2} \right) [F(t)]. \qquad \text{(VI.F-13)}$$

(The student should note that the first left hand side term in the above equation is a matrix inverse!) Since most simulations, especially those using spreadsheets, do not readily lend themselves to performing inverse matrix calculations, we shall devote a

little effort to simplifying this equation for the homogeneous force case.

The matrix to be inverted is

$$[1] + \frac{[\zeta(t)]\,\Delta t}{2} = \begin{pmatrix} 1 & \dfrac{\alpha(t+\Delta t)\,\Delta t}{2} \\ \dfrac{\beta(t+\Delta t)\,\Delta t}{2} & 1 \end{pmatrix}. \qquad \text{(VI.F-14)}$$

Inspection allows us to write the inverse as

$$\left([1] + \frac{[\zeta(t)]\,\Delta t}{2}\right)^{-1} = \begin{pmatrix} 1 & -\dfrac{\alpha(t+\Delta t)\,\Delta t}{2} \\ -\dfrac{\beta(t+\Delta t)\,\Delta t}{2} & 1 \end{pmatrix} D^{-1} \qquad \text{(VI.F-15)}$$
$$= \left([1] - \frac{\zeta(t+\Delta t)]\,\Delta t}{2} \right) D^{-1}$$

where:

$$D = \left(1 - \frac{\alpha(t+\Delta t)\ \beta(t+\Delta t)\ \Delta t^2}{4} \right). \qquad \text{(VI.F-16)}$$

The student can readily confirm that this is the inverse matrix by performing the requisite matrix multiplication.

We may now substitute this equation into equation (VI.E-13) to yield

$$[F(t+\Delta t)] \approx \left([1] - \frac{[\zeta(t+\Delta t)]\,\Delta t}{2} \right) D^{-1} \left([1] - \frac{[\zeta(t)]\,\Delta t}{2} \right) [F(t)]. \qquad \text{(VI.F-17)}$$

This equation may be multiplied out as

$$[F(t+\Delta t)] \approx \left([1] - \frac{[\zeta(t+\Delta t)]\,\Delta t}{2} - \frac{[\zeta(t)]\,\Delta t}{2} + \frac{[\zeta(t+\Delta t)]\,[\zeta(t)]\,\Delta t^2}{4} \right) D^{-1}\,[F(t)] \tag{VI.F-18}$$

$$= \begin{pmatrix} 1 - \dfrac{\alpha(t+\Delta t)\ \beta(t)\Delta t^2}{2} & -\dfrac{(\alpha(t+\Delta t) + \alpha(t)\ )\ \Delta t}{2} \\ -\dfrac{(\beta(t+\Delta t) + \beta(t)\ )\ \Delta t}{2} & 1 - \dfrac{\alpha(t)\ \beta(t+\Delta t)\ \Delta t}{4} \end{pmatrix} D^{-1}\,[F(t)].$$

The individual elements are

$$A(t+\Delta t) \approx \frac{1 - \dfrac{\alpha(t+\Delta t)\ \beta(t)\ \Delta t^2}{4}}{1 - \dfrac{\alpha(t+\Delta t)\ \beta(t+\Delta t)\ \Delta t^2}{4}} A(t) - \frac{(\ \alpha(t+\Delta t) + \alpha(t)\ )\ \Delta t}{2\left(1 - \dfrac{\alpha(t+\Delta t)\ \beta(t+\Delta t)\ \Delta t^2}{4}\right)} B(t), \tag{VI.F-19}$$

and

$$B(t+\Delta t) \approx \frac{1 - \dfrac{\beta(t+\Delta t)\ \alpha(t)\ \Delta t^2}{4}}{1 - \dfrac{\alpha(t+\Delta t)\ \beta(t+\Delta t)\ \Delta t^2}{4}} B(t) - \frac{(\ \beta(t+\Delta t) + \beta(t)\ )\ \Delta t}{2\left(1 - \dfrac{\alpha(t+\Delta t)\ \beta(t+\Delta t)\ \Delta t^2}{4}\right)} A(t), \tag{VI.F-20}$$

These equations are readily usable in a spreadsheet simulation

The matrix linear law integral equation may also be approximated using this rule, but the result is much less satisfying. The form of this solution is

$$[F(t+\Delta t)] = [F(t)] - [F^{-1}(t+\Delta t)]^{-1}\,[\zeta(t+\Delta t)]\,[F(t+\Delta t)]\,\frac{\Delta t}{2} - [F^{-1}(t)]^{-1}\,[\zeta(t)]\,[F(t)]\frac{\Delta t}{2}, \tag{VI.F-21}$$

which may be rearranged as

$$\left([1] + \frac{[F^{-1}(t+\Delta t)]^{-1}\,[\zeta(t+\Delta t)]\,\Delta t}{2}\right)[F(t+\Delta t)] = \left([1] - \frac{[F^{-1}(t)]^{-1}\,[\zeta(t)]\,\Delta t}{2}\right)[F(t)]. \tag{VI.F-22}$$

The nonsatisfactory aspect of this approximation arises from the fact that $[F^{-1}(t+\Delta t)]^{-1}$ is itself a function of the force strengths. Specifically,

$$\begin{aligned}
&[1] + \frac{[F^{-1}(t+\Delta t)]^{-1}\,[\zeta(t+\Delta t)]\,\Delta t}{2} \\
&= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \frac{\Delta t}{2}\begin{pmatrix}A(t+\Delta t) & 0\\ 0 & B(t+\Delta t)\end{pmatrix}\begin{pmatrix}0 & \alpha(t+\Delta t)\\ \beta(t+\Delta t) & 0\end{pmatrix} \\
&= \begin{pmatrix}1 & \dfrac{\alpha(t+\Delta t)\,B(t+\Delta t)\,\Delta t}{2}\\ \dfrac{\beta(t+\Delta t)\,A(t+\Delta t)\,\Delta t}{2} & 1\end{pmatrix}.
\end{aligned} \tag{VI.F-23}$$

There are at least two simple approaches to proceeding with this approximation. We may take equation (VI.E-22) and write out the elements. By cross substitution, $B(t+\Delta t)$ and $A(t+\Delta t)$ can be eliminated from the equations for $A(t+\Delta t)$ and $B(t+\Delta t)$ which are quadratic in these variable. These resulting quadratic equations can be solved algebraically. The alternate approach is to take the inverse of the above equation,

$$\left([1] + \frac{[F^{-1}(t+\Delta t)]^{-1}\,[\zeta(t+\Delta t)]\,\Delta t}{2}\right)^{-1} = \begin{pmatrix}1 & -\dfrac{\alpha(t+\Delta t)\,B(t+\Delta t)\,\Delta t}{2}\\ -\dfrac{\beta(t+\Delta t)\,A(t+\Delta t)\,\Delta t}{2} & 1\end{pmatrix}, \tag{VI.F-24}$$

and approximate $A(t+\Delta t)$ and $B(t+\Delta t)$ with their zeroth order approximate solutions.

The resulting equation is directly solvable. Development of this equation is left as an exercise for the student.

There are, of course, other numerical approximations which may be used, even for the simple homogeneous case considered as a theme here. Further techniques will be developed in later chapters as they become necessary. Key here is that we have laid the basis for the expositions on historical insights and heterogeneous forces to be covered in subsequent chapters.

1. Orvis, William J., **1-2-3 for Scientists and Engineers**, Sybex, San Francisco, 1987.

2. Carnahan, Brice, H. A. Luther, and James O. Wilkes, **Applied Numerical Methods**, John Wiley and Sons, New York, 1969.

VI.G. A Combined Law Attrition Example

In this section, we consider a set of combined law attrition differential equations,

$$\frac{dA}{dt} = -\alpha\ B - \phi\ A\ B\ , \tag{VI.G-1}$$

and

$$\frac{dB}{dt} = -\beta\ A - \psi\ A\ B, \tag{VI.G-2}$$

where the attrition of each force is both linear (area fire) and quadratic (direct fire.) The Red and Blue forces are considered to be comprised of units which both use weapons which have area and direct fire characteristics. This problem verges on a heterogeneous force situation (where there would be direct fire forces and area fire forces which all attrit each other,) but makes no distinction as to any differential losses between the two types of weapons, nor any distinction as to the composition of the units or forces. Any fire allocation fractions (the relative amount of each force allocated to each of the two types of fire) is assumed to be included in the attrition rates, which, for convenience, are also taken to be constants. An example of this type of force could be one comprised entirely of infantry units which completely integrate rifles (direct fire weapons) and machine guns (which have a beaten zone of fire - area weapons,) and attrition is such that no distinction in losses of the two types of weapons is made. This might occur if doctrine dictates that a constant ratio of rifles to machine guns is maintained in each unit as it takes losses. Soldiers in the units would then be presumed to be crosstrained in the use of both types of weapons and ammunition resupply would not be a problem.

If we factor equations (VI.G-1) and (VI.G-2), and eliminate time as the independent variable, the single differential equation in the force strengths is

$$\frac{dA}{dB} = \frac{(\alpha\ +\ \phi A)\ B}{(\beta\ +\ \psi B)\ A}, \tag{VI.G-3}$$

which may be written in its exact form as

$$\frac{A\ dA}{\alpha\ +\ \phi A} = \frac{B\ dB}{\beta\ +\ \psi}\ . \tag{VI.G-4}$$

Integration of this exact differential equation gives the state solution as

$$\frac{A - A_0}{\phi} - \frac{\alpha}{\phi^2} \ln\left(\frac{\alpha + \phi A}{\alpha + \phi A_0}\right) = \frac{B - B_0}{\psi} - \frac{\beta}{\psi^2} \ln\left(\frac{\beta + \psi B}{\beta + \psi B_0}\right). \qquad \text{(VI.G-5)}$$

This state solution is clearly transcendental - it cannot be written in a form which allows one force strength to be directly expressed in terms of the other force strength and the initial force strengths and attrition rates. Thus, the method of normal forms described in Chapter III cannot be applied to this problem. As such, it represents one of a class of problems which possess state solutions, but which cannot be solved using the normal forms method because of the complex form of the state solution.

If we limit the solution to small losses only, where A and B do not greatly differ from A_0 and B_0, respectively, then the state solution may be expanded. (The validity of this approximation will be examined in the subsequent chapter on historical insights.) To perform this expansion, we write

$$\begin{aligned} A &= A_0 - \Delta A \\ B &= B_0 - \Delta B, \end{aligned} \qquad \text{(VI.G-6)}$$

and rewrite the state solution as

$$\frac{\Delta A}{\phi} - \frac{\alpha}{\phi^2} \ln\left(1 - \frac{\phi \Delta A}{\alpha + \phi A_0}\right) = \frac{\Delta B}{\psi} - \frac{\beta}{\psi^2} \ln\left(1 - \frac{\psi \Delta B}{\beta + \psi B_0}\right). \qquad \text{(VI.G-7)}$$

which may now be expanded using the previously defined expansion of the logarithm. This gives

$$\frac{2\alpha + \phi A_0}{\alpha + \phi A_0} \frac{\Delta A}{\phi} = \frac{2\beta + \psi B_0}{\beta + \psi B_0} \frac{\Delta B}{\psi}. \qquad \text{(VI.G-8)}$$

This approximate state solution may be written as

$$\Delta A = \frac{\phi}{\psi} \frac{2\beta + \psi B_0}{\beta + \psi B_0} \frac{\alpha + \phi A_0}{2\alpha + \phi A_0} \Delta B. \qquad \text{(VI.G-9)}$$

If we define the constant term on the right hand side as θ, then this equation may be rewritten, using the definitions of ΔA and ΔB, as

$$A = \theta B + (A_0 - \theta B_0), \qquad \text{(VI.G-10)}$$

a linear-like approximation to the state solution which is actually infinitely multiple order (linear plus quadratic plus all higher orders.)

If equation (VI.G-10) is substituted into equation (VI.G-2), the resulting

homogeneous differential equation is

$$\frac{dB}{dt} = -(\beta + \psi B)\left(\theta B - (A_0 - \theta B_0)\right), \qquad \text{(VI.G-11)}$$

which has the solution

$$B(t) = \frac{(A_0 - \theta B_0)(\beta + \psi B_0)\, e^{-[\psi(A_0 - \theta B_0) - \beta\theta]t} - \beta A_0}{\psi A_0 - \theta(\beta + \psi B_0)\, e^{-[\psi(A_0 - \theta B_0) - \beta\theta]t}}. \qquad \text{(VI.G-12)}$$

The Red force solution can be derived in the same manner, but for brevity is merely presented here as

$$A(t) = \frac{\psi A_0(A_0 - \theta B_0) - \beta\theta A_0}{\psi A_0 - \theta(\beta + \psi B_0)\, e^{-[\psi(A_0 - \theta B_0) - \beta\theta]t}}. \qquad \text{(VI.G-13)}$$

Derivation of this equation is left as an exercise for the student.

Examination of equations (VI.G-12) and (VI.G-13) reveals that the solutions of the two equations are linear-like in form. This is a direct result of the approximation process.

While approximate time solutions of the force strengths could be calculated using numerical approximation, an analytical approximation technique was selected to be an example of how such approximations may be applied. In general, analytical approximations are preferred to numerical ones because of the portability of the approximations and the analytical solutions are easier to probe mathematically for insight, which is the ultimate goal of the techniques described in this book. Of course, an exact analytical solution is always the preferred result, but in many cases of greater complexity than the pure attrition differential equation forms of linear, quadratic, and mixed law, the state solution either does not exist or is transcendental (as in this example) and cannot be use in solving the differential equations. Failing thus to be able to solve the attrition differential equations exactly, the investigator is forced to make use of some method of approximation to find solutions. Both analytical and numerical approximations have their places and relative merits and demerits in the search for solutions. Approximate analytical solutions have the advantage that they are more easily manipulated on paper, and the interplay of parameters such as the attrition rates may be more clearly seen. They have the disadvantage that the approximation generally has limits on its applicability and care must be taken not to draw conclusions which are too general and transcend the limits of the approximation. Numerical approximations may actually be more generally accurate within the limits on the step size of the numerical integration, but they are hopelessly coupled (in most cases,) to the digital computer, and insight into the

interplay of parameters must be gained laboriously from repeated calculations. A good technical approach therefore is to make use of both types of approximation in working real problems, combining the insightful nature of the analytical approximations with the relative exactness of the numerical approximations.

The differential equations of this example can, of course, be solved numerically, but it should be noted that the form of the matrix equations would be different from those derived in the preceding section. Using the notation of the preceding two sections, we may write the matrix integral equations for the differential equations of this section as

$$[F(t+\Delta t)] \approx [F(t)] - \int_t^{t+\Delta t} [\zeta]\,[F(t')]\,dt' - \int_t^{t+\Delta t} [F^{-1}(t')]^{-1}\,[\eta]\,[F(t')]\,dt', \qquad \text{(VI.G-14)}$$

where: $[\zeta]$ is the array of quadratic-like attrition rates, and
$[\eta]$ is the array of linear-like attrition rates.

If the student wishes to solve this integral equation numerically, care should be taken with use of the trapezoid rule because of the complexity of the resulting solutions. Use of the rectangular rule with a smaller time step size is fraught with much less difficulty.

VI.H. Quadratic Lanchester Law with Reinforcements

In the preceding section, we considered an example of a set of attrition differential equations which had a transcendental state solution. In this section we consider a set of attrition differential equations which do not possess a state solution at all. The problem that we consider here is that of the classic quadratic law differential equations with the addition that reinforcement of the forces is included. The differential equations are

$$\frac{dA}{dt} = -\alpha\; B + a(t), \tag{VI.H-1}$$

and

$$\frac{dB}{dt} = -\beta\; A + b(t), \tag{VI.H-2}$$

where a(t) and b(t) are the reinforcement rates of the red and blue forces, respectively. For generality, these reinforcement rates are assumed to be known functions of time: they may be constants or they may be punctuated - that is, reinforcements arrive only at certain times and in numbers. It is assumed that the reinforcement rates are not functions of the force strengths - that problem will be considered in a later chapter.*

This set of attrition differential equations does not possess a state solution since a(t) and b(t) are presumed to be functions of time. In the extreme case where the reinforcement rates are constant, then a state solution does exist. That special case is not considered here as its solution is a direct application of the methods developed earlier, largely in Chapter III. (The method of normal forms may be applied.)

If we were to proceed to solve this set of attrition differential equations numerically, the matrix integral equation could be formed immediately as

$$\begin{aligned}[F(t+\Delta t)] = [F(t)] - \int_t^{t+\Delta t} [\zeta]\, [F(t')]\, dt' \\ + \int_t^{t+\Delta t} [R(t')]\, dt',\end{aligned} \tag{VI.H-3}$$

where [R(t)] is the matrix of reinforcement rates. This integral equation can be solved numerically using the integration approximation techniques described in preceding sections of this chapter. We shall not invoke those techniques in this section (except

* Care must be taken when the reinforcement rates are punctuated when using numerical approximations. In an analytical sense they are represented by a Dirac delta function so that when they are carried over, these are replaced by one (or more) Kronecker delta functions.

as a sidebar,) but we will in the conclusion to this section invoke a similar approximation.

Rather than pursue a purely numerical approximation, the approach here will be to pursue an analytical solution to the maximum extent possible. Engel[1] solved this problem for a(t) = 0 in his consideration of the Iwo Jima campaign of World War II. The derivation here parallels his technique but expands to the more general case where both forces have nonzero replacement rates.

In solving the attrition differential equations, equations (VI.H-1) and (VI.H-2), use is made of the analytic solutions of the quadratic law attrition differential equations, equations (III.A-4) and (III.A-5), solved in Chapter III, to write solutions of the form

$$A(t) = A^*(t) \cosh(\gamma t) - \delta\, B^*(t) \sinh(\gamma t), \tag{VI.H-4}$$

and

$$B(t) = B^*(t) \cosh(\gamma t) - \frac{A^*(t)}{\delta} \sinh(\gamma t). \tag{VI.H-5}$$

If we differentiate equations (VI.G-4) and (VI.G-5), and substitute both these equations and their derivatives into equations (VI.G-1) and (VI.G-2), and remove those terms represented by equations (VI.G-4) and (VI.G-5) (those undifferentiated in $A^*(t)$ and $B^*(t)$,) the resulting differential equations are

$$\cosh(\gamma t) \frac{dA^*}{dt} - \delta \sinh(\gamma t) \frac{dB^*}{dt} = a(t), \tag{VI.H-6}$$

and

$$\cosh(\gamma t) \frac{dB^*}{dt} - \frac{1}{\delta} \sinh(\gamma t) \frac{dA^*}{dt} = b(t), \tag{VI.H-7}$$

For this solution, it is convenient to write equations (VI.G-6) and (VI.G-7) in matrix form as

$$\begin{pmatrix} \cosh(\gamma t) & -\delta \sinh(\gamma t) \\ -\frac{1}{\delta} \sinh(\gamma t) & \cosh(\gamma t) \end{pmatrix} \frac{d}{dt}\begin{pmatrix} A^* \\ B^* \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}. \tag{VI.H-8}$$

Since the leading matrix on the left hand side is hyperbolic, we make use of its inverse to rewrite equation (VI.G-8) as

$$\frac{d}{dt}\begin{pmatrix} A^* \\ B^* \end{pmatrix} = \begin{pmatrix} \cosh(\gamma t) & \delta \sinh(\gamma t) \\ \frac{1}{\delta} \sinh(\gamma t) & \cosh(\gamma t) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \qquad \text{(VI.H-9)}$$

This differential equation is exact since all of the terms on the left hand side are known, and can be readily solved as

$$\begin{pmatrix} A^* \\ B^* \end{pmatrix} = \int_0^t dt' \begin{pmatrix} \cosh(\gamma t') & \delta \sinh(\gamma t') \\ \frac{1}{\delta} \sinh(\gamma t') & \cosh(\gamma t') \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} A^*_0 \\ B^*_0 \end{pmatrix}, \qquad \text{(VI.H-10)}$$

which allows us to write the exact solution, Equations (VI.G-4) and (VI.G-5), as

$$\begin{aligned} A(t) &= A_0 \cosh(\gamma t) - \delta\ B_0 \sinh(\gamma t) \\ &+ \cosh(\gamma t) \int_0^t dt' \left(a(t') \cosh(\gamma t') + \delta\ b(t') \sinh(\gamma t')\right) \qquad \text{(VI.H-11)} \\ &- \delta \sinh(\gamma t) \int_0^t dt' \left(b(t') \cosh(\gamma t') - \frac{a(t')}{\delta} \sinh(\gamma t')\right), \end{aligned}$$

and

$$\begin{aligned} B(t) &= B_0 \cosh(\gamma t) - \frac{A_0}{\delta} \sinh(\gamma t) \\ &+ \cosh(\gamma t) \int_0^t dt' \left(b(t') \cosh(\gamma t') + \frac{a(t')}{\delta} \sinh(\gamma t')\right) \qquad \text{(VI.H-12)} \\ &- \frac{1}{\delta} \sinh(\gamma t) \int_0^t dt' \left(a(t') \cosh(\gamma t') - \delta\ b(t') \sinh(\gamma t')\right). \end{aligned}$$

By using the properties of the hyperbolic functions for addition of arguments, equations (VI.G.8) may be rewritten in a more compact form as

$$\begin{aligned} A(t) &= A_0 \cosh(\gamma t) - \delta\ B_0 \sinh(\gamma t) \\ &+ \int_0^t dt'\ a(t') \cosh(\gamma t - \gamma t') \qquad \text{(VI.H-13)} \\ &- \delta \int_0^t dt'\ b(t') \sinh(\gamma t - \gamma t'), \end{aligned}$$

and

$$B(t) = B_0 \cosh(\gamma t) - \frac{A_0}{\delta} \sinh(\gamma t) + \int_0^t dt' \, b(t') \cosh(\gamma t - \gamma t') - \frac{1}{\delta} \int_0^t dt' \, a(t') \sinh(\gamma t - \gamma t'), \qquad \text{(VI.H-14)}$$

which clearly shows, even by inspection, that equations (VI.G-13) and (VI.G-14) are the analytical solutions of equations (VI.G-1) and (VI.G-2). If a(t) and b(t), the reinforcement rates are integrable with respect to the hyperbolic functions, equations (VI.G-13) and (VI.G-14) are readily usable in calculations.

In his 1954 paper, Engel also presents a finite difference (numerical) approximation to these equations which he uses to analyze the Iwo Jima campaign. This approximation is necessitated because force strengths and reinforcement rates are known only for whole days - the structure of the reinforcement rates is not known. As we shall address in the subsequent chapter on historical insights, this lack of precise force structure information is one of the primary difficulties in historical analysis of Lanchester's Laws. As a prelude to that chapter, we here derive a pair of finite difference approximate solutions.

The easiest way to form this approximation, rather than using equation (VI.G-3) is to start with equation (VI.G-13) (assuming that we can carry equation (VI.G-14) in parallel, but neglecting that derivation due to the cumbersome nature of the algebra,) written as

$$A(t+\Delta t) = A_0 \cosh(\gamma t+\gamma \Delta t) - \delta B_0 \sinh(\gamma t+\gamma \Delta t) + \int_0^{t+\Delta t} dt' \, a(t') \cosh(\gamma t + \gamma \Delta t - \gamma t') - \delta \int_0^{t+\Delta t} dt' \, b(t') \sinh(\gamma t + \gamma \Delta t - \gamma t'), \qquad \text{(VI.H-15)}$$

which we expand to factor out terms in the hyperbolic functions of $\gamma\Delta t$. We next split the integrals into two parts, one over the interval 0 to t, and the other over the interval t to $t+\Delta t$, and collect terms. This allows us to clearly recognize the two terms that are just A(t) and B(t), equations (VI.G-13) and (VI.G-14), times $\cosh(\gamma\Delta t)$ and $\sinh(\gamma\Delta t)$, respectively. If we replace these terms, this becomes

$$A(t+\Delta t) = A(t)\cosh(\gamma\Delta t) - \delta\, B(t)\sinh(\gamma\Delta t)$$
$$+ \int_t^{t+\Delta t} dt'\, a(t')\cosh(\gamma t + \gamma\Delta t - \gamma t') \qquad \text{(VI.H-16)}$$
$$- \delta \int_t^{t+\Delta t} dt'\, b(t')\sinh(\gamma t + \gamma\Delta t - \gamma t').$$

The last analytical step is to change the variable of integration in the two integrals on the left hand side of this equation to have limits between 0 and Δt. This gives

$$A(t+\Delta t) = A(t)\cosh(\gamma\Delta t) - \delta\, B(t)\sinh(\gamma\Delta t)$$
$$+ \int_0^{\Delta t} dt\, a(t+t')\cosh(\gamma\Delta t - \gamma t') \qquad \text{(VI.H-17)}$$
$$- \delta \int_0^{\Delta t} dt'\, b(t+t')\sinh(\gamma\Delta t - \gamma t').$$

This equation has the form of equation (VI.G-13).

To this point, the solution is exact and analytic; it incorporates the finite difference in time without approximation. To proceed however, we now introduce two different approximations. The first of these is to apply the rectangular rule (described in an earlier section,) to obtain the result

$$A(t+\Delta t) = A(t)\cosh(\gamma\Delta t) - \delta\, B(t)\sinh(\gamma\Delta t) \qquad \text{(VI.H-18)}$$
$$+ a(t)\cosh(\gamma\Delta t) - \delta\, b(t)\sinh(\gamma\Delta t),$$

while the corresponding Blue force strength approximation is

$$B(t+\Delta t) = B(t)\cosh(\gamma\Delta t) - \frac{A(t)}{\delta}\sinh(\gamma\Delta t)$$
$$+ b(t)\cosh(\gamma\Delta t) - \frac{a(t)}{\delta}\sinh(\gamma\Delta t). \qquad \text{(VI.H-19)}$$

We note that we could equally well have applied the trapezoid rule to achieve a somewhat more accurate approximation. That development is left as an exercise for the student.

An alternate approximation (that used by Engel) is to treat a(t) and b(t) as constant over the interval t to t + Δt with values of the lower limit. This allows the integrals to be performed analytically with the result,

$$A(t+\Delta t) = A(t)\cosh(\gamma\Delta t) - \delta\, B(t)\sinh(\gamma\Delta t)$$
$$+ \frac{a(t)\sinh(\gamma\Delta t)}{\gamma} - \frac{b(t)\,(\cosh(\gamma\Delta t) - 1)}{\beta}. \qquad \text{(VI.H-20)}$$

The corresponding Blue force strength approximation is

$$B(t+\Delta t) = B(t)\cosh(\gamma\Delta t) - \frac{A(t)}{\delta}\sinh(\gamma\Delta t) + \frac{b(t)\sinh(\gamma\Delta t)}{\gamma} - \frac{a(t)\,(\cosh(\gamma\Delta t) - 1)}{\alpha}. \tag{VI.H-21}$$

These equations are exactly the form presented by Engel except that there is only one replacement rate in his problem.

The student will note that an alternate approach to Engel's problem would have been to use a bilinear approximation for the attrition rates, thus incorporating both upper and lower limit values of these rates. This approximation would have been slightly more accurate, but was not used. The technique used here does, however, illustrate an analytical approximation approach - the approximation of functions under the integral. We will use this technique again in later problems.

Although it is not necessarily clear at this point, the $\cosh(\gamma\Delta t)/\sinh(\gamma\Delta t)$ approximation is also more accurate than the use of the rectangular and trapezoid rules. Derivation of this approximation arises from the repeated resubstitution of the integral equations that was mentioned in an earlier section. We will revisit this approximation in a later chapter when we consider explicitly timedependent attrition rates.

1. Engel, J. H., "A Verification of Lanchester's Law", Operations Research, 2 163, 1954.

VII. MATHEMATICAL THEORY III: Solutions of the Osipovian Attrition Differential Equations

VII.A. Introduction

In Chapter II, we briefly reviewed the attrition theory of Osipov, in which we introduced the derivation of the (Lanchester-Osipov) Quadratic Law analytic solution, the analysis of historical data, and the 3/2 attribution law. In this chapter, we return to the study of Osipovian attrition by examining the other solutions of the Lanchester-Osipov attrition differential equations. Osipov himself was primarily concerned with the analysis of historical data, and dismissed the need to investigate the time solutions of the force numbers.[a] Given the computational restrictions of the period, this is readily understandable. In today's period of easy numerical calculation by electronic means, the time solutions are useful.

The general form of the attrition differential equations are:

$$\frac{dA}{dt} = -\alpha\, A^{2-n}\, B, \qquad \text{(VII.A-1)}$$

and

$$\frac{dB}{dt} = -\beta\, B^{2-n}\, A, \qquad \text{(VII.A-2)}$$

where: A, B are the number of units of the red, blue forces,
α, ß are the attrition rate constants/functions of the two forces,
t is time, and
n is the attrition order.

We have seen previously that if n = 1, the resulting attrition differential equations give rise to what we traditionally know as Lanchester's Linear Law (and its associated solutions), and if n = 2, the attrition differential equations give rise to what we traditionally know as Lanchester's Quadratic Law (and its associated solutions) which we now know was also derived by Osipov and which was originally solved by him. If we were scrupulously correct, we should likely refer to this law as the Lanchester-Osipov Quadratic Law (using alphabetic order of the names) or even Chase-Lanchester-Osipov, but tradition in the literature has accustomed us to thinking of the law only as Lanchester's. In the interest of brevity, this connection will be maintained

[a] Although he did publish the time solutions of the quadratic equations.

in this book. Unfortunately, this does not do justice to Osipov whose contribution is at least comparable to Lanchester's. Rather than restructure what we have come to think of the nomenclature of attrition theory, we shall adopt a new nomenclature based on the attrition order. If n has exactly integer values (normally one or two), we will continue to use the existing Lanchester nomenclature. If it is non integer, we will refer to the theory as Osipovian. If the value of the attrition order is unspecified, or the result is general, we will use the term Lanchester-Osipov whenever the usage is not too cumbersome or clarity is required. We will reserve reference to Chase or Fiske to the equation forms which bear close resemblance to those they developed.

VII.B. The Lanchester-Osipov State Solution

The Lanchester - Osipov (or just LO for brevity's sake), state solution may be derived in the same manner as before. We may take the attrition differential equations, Equations (VII.A-1) and (VII.A-2), and ratio then to form

$$\frac{dA}{dB} = \frac{\alpha\ A^{2-n}\ B}{\beta\ B^{2-n}\ A}, \tag{VII.B-1}$$

rewrite the fraction on the right hand side as

$$\frac{dA}{dB} = \frac{\alpha\ B^{n-1}}{\beta\ A^{n-1}}, \tag{VII.B-2}$$

and form the exact differential equation

$$\beta\ A^{n-1}\ dA = \alpha\ B^{n-1}\ dB, \tag{VII.B-3}$$

which we may solve by direct integration as

$$\frac{\beta}{n}\left(A^{n} - A_0^{n}\right) = \frac{\alpha}{n}\left(B^{n} - B_0^{n}\right). \tag{VII.B-4}$$

If we cancel the common denominators (which are just the attrition order), then the LO state solution

$$\beta\left(A^{n} - A_0^{n}\right) = \alpha\left(B^{n} - B_0^{n}\right). \tag{VII.B-5}$$

results. From a mathematical sense, we note that this integration is general and does not depend on the attrition order as long as the attrition order is not equal to zero. A state solution still exists if the attrition order is zero, but it has a different form. (We shall consider the special case of zero attrition order in a later chapter.)

The LO state solution is generally true for any nonzero value of the attrition order. For integral values of the attrition order (normally one or two), this state solution becomes what we normally think of as the Lanchester Linear and Quadratic Law State Solutions, respectfully. If the attrition order has a value of 3/2, then this state solution becomes the Osipov 3/2 Law State Solution.

VII.C. Time Solution of the Osipov Attrition Differential Equations.

In this section, the time solution of the Osipov attrition differential equations is derived. In terms of our new nomenclature, this restricts the attrition order to take on noninteger values. The techniques described in Chapter III, especially the powerful method of normal forms, do not apply because, in general, the required integrals do not exist. This may easily be seen if we rewrite the Equation (VII.B-5) state solution in the form,

$$A^n = \frac{\alpha}{\beta} B^n + \frac{\Delta_n}{\beta}, \qquad \text{(VII.C-1)}$$

where

$$\Delta_n \equiv \beta A_0^n - \alpha B_0^n, \qquad \text{(VII.C-2)}$$

is the general conclusion condition. Equation (VII.C-1) may be rewritten as

$$A = \left(\frac{\alpha}{\beta} B^n + \frac{\Delta_n}{\beta} \right)^{\frac{1}{n}}. \qquad \text{(VII.C-3)}$$

If we substitute this into the blue force attrition differential equation, Equation (VII.A-1) we get

$$\frac{dB}{dt} = -\beta B^{2-n} \left(\frac{\alpha}{\beta} B^n + \frac{\Delta_n}{\beta} \right)^{\frac{1}{n}}, \qquad \text{(VII.C-4)}$$

which may be difficult to solve even if the attrition order is an integer, fundamentally being dependent on performing the B integration. For $n = 1$, the Linear Law case, and $n = 2$, the Quadratic Law case, these integrals exist, as we have seen. For n general in value (non-integer), this is not seem generally possible.

Instead of proceeding in this manner, we make note of the differential properties of two functions of the force numbers. We define

$$f_1 \equiv \beta A^n + \alpha B^n, \qquad \text{(VII.C-5)}$$
$$f_2 \equiv A B,$$

which have derivatives
and

$$\frac{df_1}{dt} = n \beta A^{n-1} \frac{dA}{dt} + n \alpha B^{n-1} \frac{dB}{dt} \qquad \text{(VII.C-6)}$$
$$= -2 n \alpha \beta f_2,$$

$$\frac{df_2}{dt} = \frac{dA}{dt} B + A \frac{dB}{dt} \qquad \text{(VII.C-7)}$$
$$= -f_2^{\,2-n} f_1.$$

Equation (VII.C-7) may be rewritten as

$$\frac{df_2^{\,n-1}}{dt} = -(n - 1) f_1. \qquad \text{(VII.C-8)}$$

If we differentiate this equation again, and substitute Equation (VII.C-6) into it, we get,

$$\frac{d^2 f_2^{\,n-1}}{dt^2} = -(n - 1) \frac{df_1}{dt} \qquad \text{(VII.C-9)}$$
$$= 2 n (n - 1) \alpha \beta f_2.$$

This nonlinear second order differential equation in f_2 is daunting, but is directly solvable. We make the substitution

$$f_2 = f_2(0) g^l, \qquad \text{(VII.C-10)}$$

which gives us a very complicated equation

$$f_2(0)^{n-1} \left(l(n - 1) [l(n - 1) - 1] g^{l(n - 1) - 2} \left(\frac{dg}{dt}\right)^2 + l(n - 1) g^{l(n - 1) - 1} \frac{d^2 g}{dt^2} \right) \qquad \text{(VII.C-11)}$$
$$= 2 n (n - 1) \gamma^2 f_2(0) g^l,$$

where we have replace $\alpha \beta$ with γ^2. We rearrange this into
This still appears horribly difficult, but if we now select l to satisfy

$$l + 2 - l (n - 1) = 0, \qquad \text{(VII.C-13)}$$

$$l(n-1)\,[l(n-1)-1]\left(\frac{dg}{dt}\right)^2 + l(n-1)\,g\,\frac{d^2g}{dt^2}$$
$$= 2\,n\,(n-1)\,\gamma^2\,f_2(0)^{2-n}\,g^{l+2-l(n-1)}. \qquad \text{(VII.C-12)}$$

that is,

$$l = \frac{2}{n-2}, \qquad \text{(VII.C-14)}$$

this reduces to

$$\frac{2\,n\,(n-1)}{(n-2)^2}\left(\frac{dg}{dt}\right)^2 + \frac{2\,(n-1)}{n-2}\,g\,\frac{d^2g}{dt^2} = 2\,n\,(n-1)\,\gamma^2\,f_2(0)^{2-n}. \qquad \text{(VII.C-15)}$$

We note that since we expect $n \leq 2$ as a rule, Equations (VII.C-10) will have the form

$$f_2(t) = \frac{f_2(0)}{g(t)^{\frac{2}{2-n}}}, \qquad \text{(VII.C-16)}$$

so that g will increase with time. Recognizing this, we may comfortably rewrite Equation (VII.C-15) in a somewhat simpler form,

$$n\left(\frac{dg}{dt}\right)^2 - (2-n)\,g\,\frac{d^2g}{dt^2} = n\,(2-n)^2\,\gamma^2\,f_2(0)^{2-n}. \qquad \text{(VII.C-17)}$$

and apply a series solution of the form

$$g(t) = \sum_{j=0}^{\infty} g_j\,t^j\,. \qquad \text{(VII.C-18)}$$

Since Equation (VII.C-17) contains terms of order g^2, we must have two indices of summation. Thus, on substitution of Equation (VII.C-18) into Equation (VII.C-17) we get

$$n \sum_{j,k=0}^{\infty} (j + 1)(k + 1)\, g_{j+1}\, g_{k+1}\, t^{j+k}$$
$$- (2 - n) \sum_{j,k=0}^{\infty} (k + 1)(k + 2)\, g_j\, g_{k+2}\, t^{j+k} \qquad \text{(VII.C-19)}$$
$$= n\,(2 - n)^2\, \gamma^2\, f_2(0)^{2-n} .$$

This equation is still daunting, but we can simplify it still further by a simple trick - we note that the time t is always raised to the same power which is the sum of the indices j and k. We may therefore redefine our double summations with a new index of summation l (different from the previous one in this section!,) that is just the sum of j and k, and redefine one of the other indices, say k, so that it now runs only over values from zero to l. The student may wish to verify that by making this change of index, we neither introduce nor eliminate any terms. If we make this change, Equation (VII.C-19) becomes

$$\sum_{l=0}^{\infty} t^l \sum_{k=0}^{l} n\,(l - k + 1)(k + 1)\, g_{l-k+1}\, g_{k+1}$$
$$- (2 - n)(k + 1)(k + 2)\, g_{l-k}\, g_{k+2} \qquad \text{(VII.C-20)}$$
$$= n\,(2 - n)^2\, \gamma^2\, f_2(0)^{2-n} .$$

We note that this equation has the happy property of having all time dependence on the right hand side of the equation outside the second summation. If we assume linear independence of the summation terms, we may decompose Equation (VII.C-20) term by term, matching powers of time.

Before preceding to this however, it is useful to backtrack at this point and determine the values of the first two terms in the g expansion. Clearly, the zeroth index term is one from Equation (VII.C-10). Since this is the only term in the g expansion that contributes to the value of g at t = 0, this assignment is necessary to assure that f_2 has the proper value at t = 0. The first index term has a value given by Equation (VII.C-7), evaluated at t = 0. This action is just the application of the second (necessary) boundary condition. Thus, the first two terms in the g expansion have values:

$$g_0 = 1 ,$$
$$g_1 = \frac{2 - n}{2}\, f_2(0)^{1-n}\, f_1(0) . \qquad \text{(VII.C-21)}$$

$$n\, g_1^2 - 2\,(2-n)\, g_0\, g_2 = n\,(2-n)^2\, \gamma^2\, f_2(0)^{2-n}\,. \qquad \text{(VII.C-22)}$$

We may now proceed to the decomposition of Equation (VII.C-20). For $l = 0$, the decomposition has the simple form, which, since $g_0 = 1$, has the simple form,

$$g_2 = \frac{n\,(2-n)}{8}\, f_2(0)^{2-2n}\,(f_1(0)^2 - 4\,\gamma^2\, f_2(0)^n)\,. \qquad \text{(VII.C-23)}$$

We note that the last term of Equation (VII.C-23) is just the state solution in terms of f_1 and f_2. Further, the constant left hand side of Equation (VII.C-20) only contributes to the first, $l = 0$, decomposition term. Higher order in l decomposition terms ($l > 0$), may be generally written as,

$$g_{l+2} = \frac{\frac{n}{2-n}\sum_{k=0}^{l}(l-k+1)\,(k+1)\, g_{l-k+1}\, g_{k+1} - \sum_{k=0}^{l-1}(k+1)\,(k+2)\, g_{l-k}\, g_{k+2}}{(l+1)\,(l+2)}\,, \qquad \text{(VII.C-24)}$$

where we have used $g_0 = 1$. This equation allows us to calculate the terms of the g expansion, starting with g_0, g_1, and g_2, in a bootstrap fashion.

VII.D Near Quadratic Behavior of the Osipov Time Solution

Clearly, with $n = 2$, the mathematics of the preceding section is unnecessary, we may solve the attrition differential equations directly by the method of normal forms. It is valuable, however, to examine the behavior of our formalism when $n = 2$.

We may expand f_2^{n+1} about $n = 2$ by rewriting it as an exponential and expanding,

$$\begin{aligned} f_2^{n-1} &= e^{(n-1)\ln(f_2)}\,, \\ &= f_2 \sum_{j=0}^{\infty} \frac{\ln(f_2)^j}{j!}\,(n-2)^j\,. \end{aligned} \qquad \text{(VII.D-1)}$$

The first order term is then just

$$f_2^{n-1} = f_2\,[\,1 + \ln(f_2)\,(\,n-2\,)\,]\,, \qquad \text{(VII.D-2)}$$

which has a first derivative

$$\frac{df_2^{n-1}}{dt} = \frac{df_2}{dt} [1 + \ln(f_2) (n - 2)] + \frac{df_2}{dt} (n - 2) , \qquad \text{(VII.D-3)}$$

and a second derivative

$$\frac{d^2 f_2^{n-1}}{dt^2} = \frac{d^2 f_2}{dt^2} [1 + \ln(f_2) (n - 2)] + (\frac{df_2}{dt})^2 \frac{(n - 2)}{f_2} + \frac{d^2 f_2}{dt^2} (n - 2) , \qquad \text{(VII.D-4)}$$
$$= 2 n (n - 1) \gamma^2 f_2 .$$

and associated second order differential equation.

This is a much more complicated (because of the $\ln(f_2)$ term,) non-linear differential equation than we had before, but it is, in principle, solvable. Because this solution presents great difficulty and adds little to our exposition here that the formalism is extensible all the way to n = 2, we do not present that solution here. Rest assured however, that the mathematics continues to hold!

It is possible to form approximate solutions to this differential equation, but we shall only indicate the approach here. Since the term $\ln(f_2)$ is relatively small and changes slowly compared to f_2, we can structure approximations based on using the initial value of f_2. These tend to be valid only for small losses - the initial part of the engagement. In this case, Equation (VII.D-4) has exponential solutions with exponents ψ given by

$$\psi = \pm \sqrt{\frac{2 n (n - 1)}{(2n - 3) + \ln(f_2(0)) (n - 2)}} \gamma . \qquad \text{(VII.D-5)}$$

Solutions can then be constructed with these exponents.

VI.E. The 3/2 Law

Osipov has noted that for forces larger than 75,000 in strength, the 3/2 law seems to agree better with historical data than does the Quadratic Law. In this section, we shall examine the behavior of the Osipov time solutions for n = 3/2.

For n = 3/2, Equation (VII.C-9) becomes

$$\frac{d^2 f_2^{\frac{1}{2}}}{dt^2} = \frac{3}{2} \gamma^2 f_2 , \qquad \text{(VII.E-1)}$$

which has a solution that is an elliptic integral. Since the methodology advanced in section C of this chapter is still applicable, we shall not overwhelm the student with an exposition of elliptic integrals, leaving that to inspection of other texts.

We shall examine the properties of the 3/2 Law in the next chapter when we present calculations.

VII F Duration of the Conflict.

One of the popular, and misused applications of attrition theory is to calculate the time to battle's end based on battle to a conclusion. This is especially true in the method used here to solve the general Osipov attrition differential equations. The f_1 function only goes to zero when both the red and blue forces have zero strength - only in the draw case. Alternately, the f_2 function goes to zero when either the red or the blue force has zero strength.

Clearly, it should be possible to extract the time t when the g function comes arbitrarily close to infinity (arbitrarily large). This value of time would correspond to a conclusion time within the definition used to determine that time. If the student has some inclination at this point that this is an entirely arbitrary definition then that is the case. The complexity of extracting the conclusion time from the g expansion (which incidentally, is no more difficult than extracting inverse tanh's,) for an arbitrarily defined value of g associated with some value of f_2 merely points up the lack of utility of the concept. The concept of conclusion time has good mathematical meaning, if properly defined in a mathematical sense. It does not appear to have good meaning in the sense of the end of a battle.

VII G Draw Solution.

The ideal solution to deal with is the true draw case - that of a complete conclusion, $\Delta_n = 0$. In this case, the state solution reduces to

$$\alpha B^n = \beta A^n , \qquad \text{(VII.G-1)}$$

which we may substitute into the attrition differential equations, Equations (VII.A-1) and (VII.A-2). This gives

$$\frac{dA}{dt} = -\alpha \, \phi \, A^{3-n} , \tag{VII.G-2}$$

and

$$\frac{dB}{dt} = -\frac{\beta}{\phi} \, B^{3-n} , \tag{VII.G-3}$$

where we have defined the new constant ϕ as

$$\phi = \left(\frac{\beta}{\alpha} \right)^{\frac{1}{n}} , \tag{VII.G-4}$$

for notational convenience. Since Equations (VII.G-2) and (VII.G-3) are now homogeneous, they may be integrated directly and with some simple algebraic manipulation, give solutions of the forms,

$$A(t) = \frac{A_0}{\left(1 + \alpha\phi(2 - n) \, A_0^{2-n} \, t \right)^{\frac{1}{2-n}}} , \tag{VII.G-5}$$

and

$$B(t) = \frac{B_0}{\left(1 + \frac{\beta}{\phi}(2 - n) \, B_0^{2-n} \, t \right)^{\frac{1}{2-n}}} . \tag{VII.G-6}$$

Note that the force strengths only become zero at infinite time. This, as we now recognize, is characteristic of draw cases - they have infinite conclusion times.

It is interesting to examine the asymptotic properties of these solutions as $n \to 2$. We may do this, for the A solution, Equation (VII.G-5), by rewriting it as

$$A(t) = A_0 \, e^{-\frac{\ln(1 + \alpha\phi(2-n)A_0^{2-n} t)}{2-n}} , \tag{VII.G-7}$$

which as n becomes close enough to 2 so that we may expand the logarithm and only keep the first order term, has the form,

$$A(t) = A_0\, e^{-\alpha\left(\frac{\beta}{\alpha}\right)^{\frac{1}{n}} A_0^{2-n}\, t}, \qquad \text{(VII.G-8)}$$

where we have explicitly written ϕ. As n becomes exactly 2, Equation (VII-G-8) becomes

$$A(t) = A_0\, e^{-\gamma t}, \qquad \text{(VII.G-9)}$$

which is the expected (desired) result.

Similar results can be obtained for the B, Equation (VII.G-6). This is left as an exercise.

VII.H. Conclusion

In this chapter, we have presented a solution method for the general Osipov- Lanchester attrition differential equations when both forces have the same attrition order n. Before closing out the chapter however, a word of warning is in order for the student who seeks to use the results of this chapter in simulation - to get numbers. There is a potential numerical instability in the form of the solution that needs accounting.

The recursion solution for the g function expansion contains a term proportional to (n - 2). This term can become numerically unstable when n is sufficiently close to 2 and recursion is made. Accordingly, care should be taken in calculations to ensure that proper numerical techniques are used.

VIII. Osipovian Attrition
Ironman Analysis and Solution Forms

VIII.A. Introduction

In this chapter we take up the analysis of Osipovian attrition using the mathematical solutions developed in the previous chapter. As in Chapter V, we again take up the pursuit of Ironman Analysis to bring some further understanding to attrition mechanics. Additionally, we examine the nature of the solution by examining specific calculations, and give some attention to the nature of the attrition processes.

VIII.B. Osipov's 3/2 Law Iron Man Analysis.

For the deterministic Ironman Analysis of Osipov's 3/2 Law, we have the differential equations

$$\frac{dA}{dt} = -\alpha \sqrt{A}, \qquad \text{(VIII.B-1)}$$

and

$$\frac{dB}{dt} = 0, \qquad \text{(VIII.B-2)}$$

since we assume the Blue force to be comprised of one Ironman (who, by assumption; cannot be attrited).

We may integrate Equation (VIII.B-1) directly since it is an exact differential equation, giving us the solution

$$\sqrt{A} = \sqrt{A_0} - \frac{\alpha\, t}{2}, \qquad \text{(VIII.B-3)}$$

where the Blue force strength, by assumption; is one. It is more convenient to write this equation in the form

$$2\left(\sqrt{A_0} - \sqrt{A}\right) = \alpha\, t, \qquad \text{(VIII.B-4)}$$

for analyzing the meaning of the attrition rate.

One natural measure of the attrition rate immediately presents itself: Let us specify that t' is the time required for the Blue force (of one unit) to reduce the Red force strength from A_0 to $A_0/4$. If we apply these conditions to equation (VIII.B-4),

force strength from A_0 to $A_0/4$. If we apply these conditions to equation (VIII.B-4), the result is

$$\sqrt{A_0} = \alpha' t', \tag{VIII.B-5}$$

or

$$\alpha' = \frac{\sqrt{A_0}}{t'}, \tag{VIII.B-6}$$

that is, α' is the square root of the initial Red force strength divided by the time that it takes the Blue force (of one!) to reduce the Red force strength by $3/4^{ths}$. This is a mathematically natural way to define the attrition rate, but it is also a bit difficult to accept on a military basis - surely no force, incapable of inflicting losses on the other side, will allow a battle to proceed to the point of loosing $3/4^{ths}$ of its strength except under the most unusual of circumstances such as combat to a conclusion or where A_0 is very small. This is surely contrary to anything Osipov would conceive of in view of his insights into historical battles. Further, this definition strains the calculability of the attrition rate - there are far too many attrition options to be resolved in one unit achieving this number of kills.

The student may, at this point, raise the question, "what of the Linear Law case?", and that question is valid. In the Linear Law Ironman Analysis (Section V.B), the attrition time τ could have been defined as the period of time required to reduce A_0 to e^{-1} of its initial value. This however, is a reduction of 66% (2/3), which is reasonably close to a 75% reduction. By assumption, the Linear Law commonly refers to indirect fire[a] which implies that the Red force is incapable of knowing directly of its attrition effects on the Blue force. If, and we here assume that, Osipov's 3/2 Law may describe direct fire combat, then the Red force is aware of its alethality. Thus, this degree of attrition (75%) is not consistent with the definition of an attrition time for the Osipov 3/2 Law.

If instead, we advance the definition introduced for the Square Law, and used throughout Chapter V, that t" is the time required to reduce the Red force strength by one unit, then we may use equation (VIII.B-4) to write

---

[a] The Linear Law may also apply to direct fire when targets are hard to find.

$$2\sqrt{A_0}\left(1 - \sqrt{1 - \frac{1}{A_0}}\right) = \alpha'' t'', \tag{VIII.B-7}$$

which if we expand the radical, assuming $A_0 \gg 1$, gives

$$\alpha'' = \frac{1}{\sqrt{A_0}\, t''}, \tag{VIII.B-8}$$

which gives an attrition rate α'' exactly A_0 times smaller than α' and $\sqrt{A_0}$ times smaller that the square law attrition rate. While this proposed definition is better than the first one, it is not satisfactory since it implies a rate of loss which is significantly less than that seen in the Square Law. We further note that if t' is linear in the individual attrition times, that is,

$$t' = \frac{A_0 t''}{4}, \tag{VIII.B-9}$$

then α' and α'' differ by a factor of 2.

There is, of course, nothing inherently wrong with the suggestion of a different time scale for Osipov's 3/2 Law, even based on the assumption of common attrition mechanisms between the 3/2 and Square Laws. It is, however, an excessive complication which we do not need to incorporate now because it implies some discontinuity in assumptions between the Linear and Square Laws (since the 3/2 Law lies between them in attrition order) which we cannot justify based on what we have examined thus far. To resolve some of this question, we now turn to the general form of Osipovian attrition.

VIII.C. General Osipovian Ironman Analysis

For the case of general attrition order, the Red force Ironman Analysis differential equation is

$$\frac{dA}{dt} = -\alpha\, A^{2-n}, \tag{VIII.C-1}$$

which has solutions

$$A_0^{n-1} - A^{n-1} = \alpha\, t\, (n - 1), \tag{VIII.C-2}$$

where $n \neq 1$.

If we again try the Square Law attrition time definition, that t" is the time required to attrit one unit, then the attrition rate is defined by

$$\alpha'' = \frac{A_0^{n-2}}{t''}, \tag{VIII.C-3}$$

which for the 3/2 Law gives

$$\alpha'' = \frac{A_0^{\frac{1}{2}}}{t''}, \tag{VIII.C-4}$$

which is identical to Equation (VIII.B-6). For the Square Law, we get

$$\alpha'' = \frac{1}{t''}, \tag{VIII.C-5}$$

which is the definition. While this proposed definition does give consistency, we will search farther along one more avenue of investigation before unconditionally adopting it.

Let the attrition time τ be the time required to attrit x units of the Red force with the restriction that x is small so that we may retain the freedom to expand the force strength terms. Thus

$$A = A_0 - x.$$

If we substitute this into equation (VIII.C-2), we get

$$A_0^{n-1} - (A_0 - x)^{n-1} = \alpha\,\tau\,(n-1). \tag{VIII.C-7}$$

which we again expand and form the equation

$$x\,A_0^{n-2} = \alpha\,\tau. \tag{VIII.C-8}$$

If we let $x = A_0^{2-n}$, then $\alpha = \tau^{-1}$, which is Bonder's equation. Now, τ is the time required to attrit A_0^{2-n} of the Red force. For the Square Law this is just one unit!, while for the 3/2 Law (and in general,) this time is force strength dependent. This does not violate any of the assumptions that we have introduced so far, but it does have the effect of introducing a degree of freedom which we must consider, that the attrition time is dependent on the force strength![a] Unfortunately, it seems to contradict the results we derived in Chapter V. We will examine this in the next section.

---

[a] We do note in passing that if τ is linear, then this result is equivalent to equation (VIII.B-8). This result should not surprise the student since the assumptions used in forming the two equations are mathematically equivalent.

Further, recall that for the Linear Law case, n = 1, we have Ironman Solutions of the form

$$\ln(A_0/A) = \alpha\tau.$$

If τ is the time that it takes to attrit the Red force to $A_0\,e^{-1}$, then this reduces to Bonder's Equation.

VIII.D. Force Strength and Attrition

In the assumptions associated with the Linear and Square Laws, the attrition form depends on whether the density of the force is a constant or varies, and that the weapons may attrit any unit in the battle (within range of the weapons). Let us now consider the consequences of these assumptions in greater detail.

Consider a force arranged in a rectangular order. The area occupied by the force is described by dimensions of depth and width, d and w, respectively. The area occupied by the unit, a, is just the product of these two. Let the unit be arranged such that the average distance between elements along the depth of the unit is d_e and along the width is w_e. Then the number of elements (units) in the area is

$$A = \frac{d\,w}{d_e\,w_e}. \tag{VIII.D-1}$$

If we introduce the ratios

$$f \equiv \frac{d}{w}, \tag{VIII.D-2}$$

and

$$f_e \equiv \frac{d_e}{w_e}, \tag{VIII.D-3}$$

we may solve equation (VIII.D-1) for the number of units along the width (or depth) of the formation as

$$A = \frac{w^2 f}{w_e^2 f_e}, \tag{VIII.D-4}$$

or

$$\frac{w}{w_e} = \sqrt{\frac{f_e}{f} A} \equiv N_w. \tag{VIII.D-5}$$

Alternately, we may write the perimeter of the formation as

$$p = 2\,(w + d)$$
$$= 2\,w\,(1 + f) = \frac{2\,d\,(1 + f)}{f}, \qquad \text{(VIII.D-6)}$$

which is related to the area by

$$a = \frac{p^2 f}{4\,(1 + f)^2}. \qquad \text{(VIII.D-7)}$$

Since the density of the force is $\rho \equiv Aa$, we may write the force strength as

$$A = \frac{\rho\, p^2 f}{4\,(1 + f)^2}. \qquad \text{(VIII.D-8)}$$

If we now use the proposed definition of attrition time advanced in the previous section, that τ was the time required to attrit A_0^{2-n} units, then from equation (VIII.D-8) we may write

$$A^{n-1} = \left(\frac{\rho\, p^2 f}{4\,(1 + f)^2}\right)^{n-1}, \qquad \text{(VIII.D-9)}$$

and now explore what this fraction means in terms of the attrition order. If n = 3/2, then 2 - n = 1/2, and the result is

$$A^{\frac{1}{2}} = \xi\,\sqrt{\rho}\,p, \qquad \text{(VIII.D-10)}$$

where ξ is a constant. If p is a constant, then this quantity varies only with the perimeter of the formation. This implies that attrition may only occur along the perimeter of the formation.

If p is a variable, then this quantity varies with the square root of the force strength, which we see from equation (VIII.D-5) is related to the number of units along the width (or depth, and thereby the perimeter of the formation.) This implies that attrition may only occur along the perimeter of the formation.

Notice that under the assumptions of either constant or variable density, which for area fire, give rise to both Linear and Quadratic Laws, the Osipov 3/2 Law develops if attrition is limited to the edge of the formation.

For direct fire, if we assume targets are hard to find, then the Linear Law again

results, but if the targets are hard to kill, and the formation has a basis where $f = f_e$, then

$$N_w = \sqrt{A}. \qquad \text{(VIII.D-11)}$$

For direct fire, we may now see that the attrition rate may be defined by Bonder's equation when τ is the time to kill one rank of the formation or by the formalism developed in Chapter V, assuming linear time.

During most of history, battles were, ideally, fought between units arranged in rectangular formations of rank and file. This is true of the era of most of the battles which Osipov considered. Thus, Osipov's 3/2 Law may be viewed as representing combat which occurred in this manner - between rank and file formations with attrition limited to the front (or sides) of the formation.

This interpretation is acceptable for direct fire weapons against this type of formation. Troops behind the first rank are hidden from view and fire by the ranks in front of them. Since direct fire in that period was usually aimed at the formation as a whole rather than at individual troops,[1] the rearward troops were shielded and attrition occurred primarily at the perimeter of the formation.

Equally clearly, however, this interpretation cannot be simply applied to indirect fire weapons which have an area coverage. While fire which landed to one side of the formation would tend to have much the same effect as direct fire, this will not explain fire which lands in the formation. Rather than postulate some strange type of attrition mechanism, let us write the force strength as

$$A = \rho_d \, d \, \rho_w \, w, \qquad \text{(VIII.D-12)}$$

where ρ_d and ρ_w are just the inverses of d_e and w_e. Both of these quantities are proportional to $\sqrt{A}$ because of the assumption of a rank and file basis. Let us now assume that the linear density along the width of the formation is constant (a constant front to the enemy). Then, ρ_w is a constant, but ρ_d is not. Rather, it is proportional to $\sqrt{A}$. Changes in the structure of the force strength can then only occur along the depth. In other words, the Osipov 3/2 Law describes area fire against a formation when the width, depth, or perimeter of the formation is constant.

This now leads us to an examination of the general form of Osipovian attrition. For area fire, a constant areal force density results in the Linear Law, a constant linear force density results in Osipov's 3/2 law, and a completely variable density results in the Square Law. Thus, the quantity (2-n) represents the fractional power of the area that the density of the unit is held constant over. In the sense of fractals,[2,3]

$$\sigma \equiv 2\,(2 - n), \qquad \text{(VIII.D-13)}$$

σ is the fractal dimension of force density under area fire. (For n = 1, σ = 2, which has dimensions of length squared, or area. For n = 3/2, σ = 1, which is a dimension of length, while for n = 2, σ = 0, which has no dimension - a point.)

For direct fire, we have sketched that the attrition order may represent a tradeoff between acquisition time and killing time. When the units are widely enough spread that all units (in theory), can be seen and shot at, the Linear and Quadratic Laws result when one or the other of these is the most time consuming. For the 3/2 Law, the two processes occur in the same manner, since units arranged across the battlefield must only be searched for in elevation, but not in azimuth (which again goes as $\sqrt{A}$) and attrition occurs only on the edges of the formation. Does the attrition order reflect some variation in the importance magnitude of the acquisition and kill times? The answer is no! The 3/2 Law equivalence is an accident which is a feature of the way that the units are arranged and attrition occurs. In direct fire attrition dominated by kill time, the attrition order n is a complex function of the force formation and where in that formation attrition can occur. In its simplest terms,

$$\sigma' = 2\,(n - 1), \qquad \text{(VIII.D-14)}$$

is the fractional dimension over which attrition can occur.

For the Linear Law case, n = 1, σ' = 0, which we may interpret as point attrition against those points (individual elements) that have been acquired. For the 3/2 Law case, n = 3/2, σ' = 1, which we interpret as attrition along a line (the width, depth, or perimeter of the formation, as appropriate.) For the Square Law, n = 2, σ' = 2, which we interpret as attrition over the entire area of the formation.

One of the questions that we have not addressed here is the way direct fire occurs in rank and file formations. While we have shown that keeping the number of troops in the front rank constant by decreasing (selectively) the number of files (or visa versa) keeps the attrition proportional to $\sqrt{A}$, we have not made the argument that since only one rank in a formation may safely and effectively fire at once, attrition is also proportional to $\sqrt{B}$ and not to B. This argument leads to a different pair of attrition differential equations which have a linear state solution. We will discuss this pair of attrition differential equations in a later chapter on alternative attrition differential equations. By training and doctrine variations, several ranks may fire at once although this reduces the firing rate (but not necessarily the loss rate.)

1. von Pivka, Otto, **Armies of the Napoleonic Era**, Tapinger Publishing Co., New York, 1979.

2. Mandelbrot, B. B., **The Fractal Geometry of Nature**, W. H. Freeman, New York, 1983.

3. Schroder, Manfred, **Fractals, Chaos, Power Laws**, W. H. Freeman and Company, New York, 1991.

VIII.E. Assumptions of the Osipov 3/2 Law

In this section, we present the assumptions associated with Osipov's 3/2 Law. Some of these have been discussed or developed in preceding sections. As we have described earlier, most of these assumptions are associated with rank and file formations of compact density used in warfare during the approximate period from (before) Alexander the Great (the Phalanx) to roughly World War I.[1] These formations, and Osipov's 3/2 Law seem especially applicable during the early period of gunpowder warfare before rate of fire became high enough to force increased dispersal which in turn forced better training and shifted the influence of the acquisition process.

Quite simply, when rate and density of fire (because of technological improvements) made attrition too fast for a rank and file formation to survive effectively (Pickett's Charge at the Battle of Gettysburg comes to mind as an example),[2,3] the survival answer was to reduce the density of the formation and to depart from the rigid formation to take advantage of terrain for protection at the individual level. This dispersal in turn required changes in the training of the troops to operate more independently (to do target acquisition and shoot without the direct control of their commissioned and non-commissioned officers - a great liberalization of armies which leveled the class structure of the army, ended the nobility - commons split in the officer - enlisted ranks with profound political implications, and increased the officer : enlisted ratio to maintain some control.) This greater dispersal and independence in turn changed the very nature of attrition. Before, acquisition and killing were separate processes. Once acquisition ended, killing began and continued until stopped. Fire was primarily directed against multi-element formations rather than against individual elements. Thus, attrition mechanics may be viewed as shifting from 3/2 Law to Linear or Quadratic Law depending on the relative temporal importance of acquisition and engagement..

This shift may, however, be also viewed as a matter of scale as well. Battles after the American Civil War continued to this day to be fought with lines between the two forces. Despite an increase in depth of weapons' effective carry, which changed the dimensionality of attrition at the local level from a line ($n = 3/2$) to an area ($n = 1$ or $n = 2$,) battles continue to occur at interfaces between the two forces. Even today, with contemporary concepts of non-linear battlefields, independently operating Corps, and the echelonment of forces, combat still has a component that occurs along the edges of its forces' formations. The bankruptcy of the Osipovian 3/2 Law should, accordingly, not be too rapidly heralded, lest it be like Mark Twain's death.

We may now state the assumptions, in the manner of Chapter IV, that we may relate to the Osipov 3/2 Law:

1.) The two forces A (for amber or red) and B (for blue) are engaged in combat.

2.) The units of the two forces are within weapons

range of units of the other side.

3.) The attrition rates are known and constant.

4a.) Each unit is aware of the general location of enemy units but is unaware of the effect of fire.

5a.) Fire is uniformly distributed over the area occupied by enemy units.

6a.) The occupied front, or depth, or perimeter density of units remains constant, units redisperse within the area to keep this dimension constant.

or

4b.) Each unit is aware of the specific location of enemy units and the effect of fire is known.

5b.) Fire from surviving units is uniformly distributed against enemy units.

6b.) The area occupied by surviving units may contract to maintain a constant linear density of units along front, depth, or perimeter.

We note immediately that assumptions 1 - 3 are identical to those that we have stated before in Chapter IV, except that we have modified assumption 2 to be consistent with our findings in that chapter about the limitations of weapons' range. Assumptions 4a - 6a are those for area fire, modified for the restriction that a constant linear density of formation is maintained. Assumptions 4b - 6b are those for point fire, again modified for the restriction that a constant linear density is maintained. Note the "may contract" in assumption 6b. If the formation maintains a constant frontal or side density (along ranks or files,) then the formation must contract. If a constant perimeter density is maintained, then the formation cannot contract although the density may decrease to keep a rectangular array of troops or a hole may appear in the center of the formation if the density along the perimeter is maintained (this is the tactic of forming a square so often used by the British in the Napoleonic Wars.)

1. Hanson, Victor Davis, **The Western Way of War,** Alfred A. Knopf, New York, 1989.

2. Griffith, Paddy, **Battle Tactics of the Civil War,** Yale University Press, New Haven, 1989.

3. Luvaas, Jay, and COL. Harold, W. Nelson, eds., **The U. S. Army War College Guide to the Battle of Gettysburg**, Harper and Row, New York, 1986.

VIII.F. Solutions of Osipov's 3/2 Law

We may now turn our attention to the form of time solutions of the Osipov 3/2 Law by examining some particular solutions. The calculation of attrition rates will use the equations derived in Section C of this chapter; calculations of the time solutions will use the numerical technique outlined in the last section of the preceding chapter to avoid the singularity in the analytical solution. For simplicity, we shall limit ourselves to formations which are square (approximately) both in terms of distance between troops within the formation and in terms of the arrangement of the formation. As in previous examples, we will use an initial red force strength, A_0, of 100 units, and an initial Blue force strength, B_0, of 200.

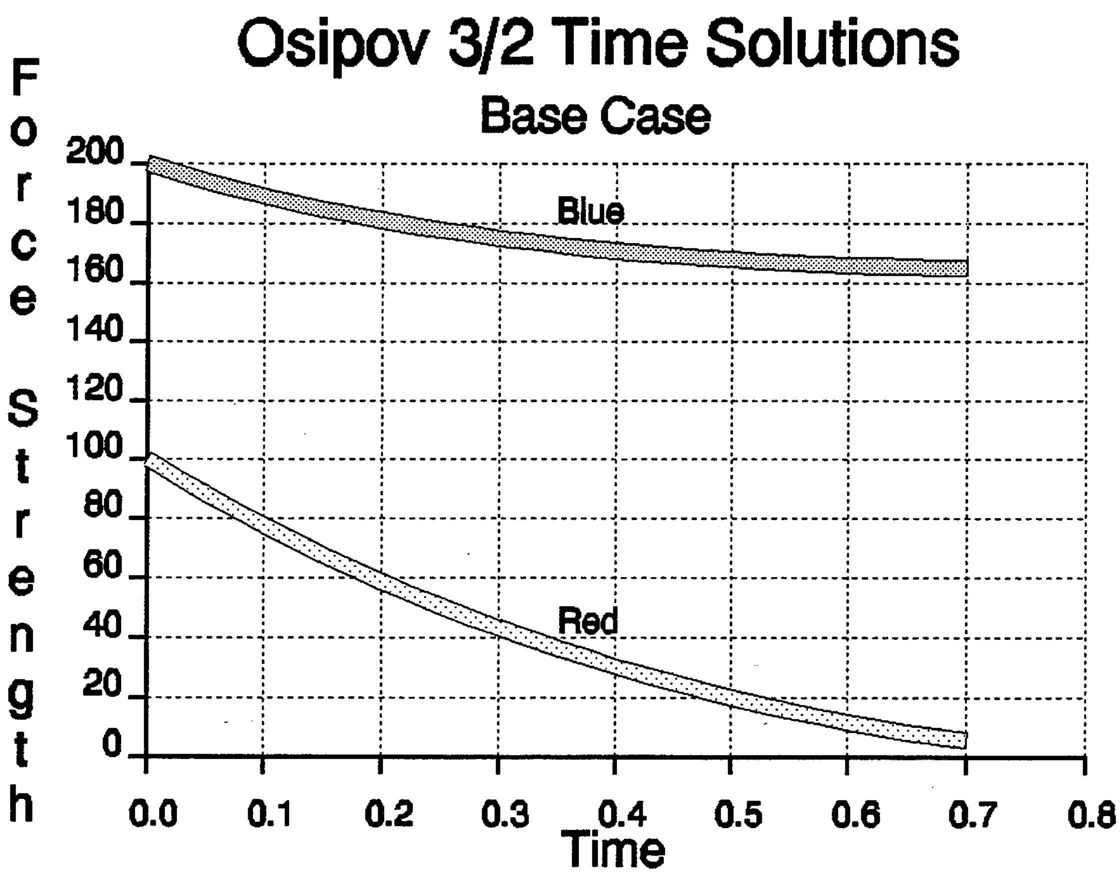


Figure (VII.F.1)

As in previous examples, we shall take the attrition time for one kill (the Square Law attrition time) to be the inverse of the product of rate of fire, r_f, and probability of kill, p_k. As an approximation, since we have not laid any basis for the attrition mechanism in a rank and file formation (again deferred to a later chapter on attrition

rate theory,) we will assume that attrition occurs linearly. Thus, the attrition time for $\sqrt{A_0}$ is just that factor times the attrition time for a single unit. For the base case, we take r_f = 4 rounds per time period, and p_k = 0.3 for both forces. We ignore the impact of acquisition time by assuming it to be small. For these parameters, the base case is shown in Figure (VIII.F.1).

As we have noted in Chapter IV, the force power of the two sides can be modified by either changing the attrition rates or the initial force strengths. In the linear law case, we saw that a change of equal magnitude of either attrition rate or initial force strength had the same effect on the outcome of the engagement, while for the square law case, we saw that it was necessary to increase the attrition rate by a factor equal to the square of the factor of increase of the initial force strength to have the same effect. This is merely the Principle of Concentration, and it demonstrated the effect of technology on the outcome of the engagement. In the Osipov 3/2 Law case, we would expect a situation intermediary between these two results just from the form of the state solution.

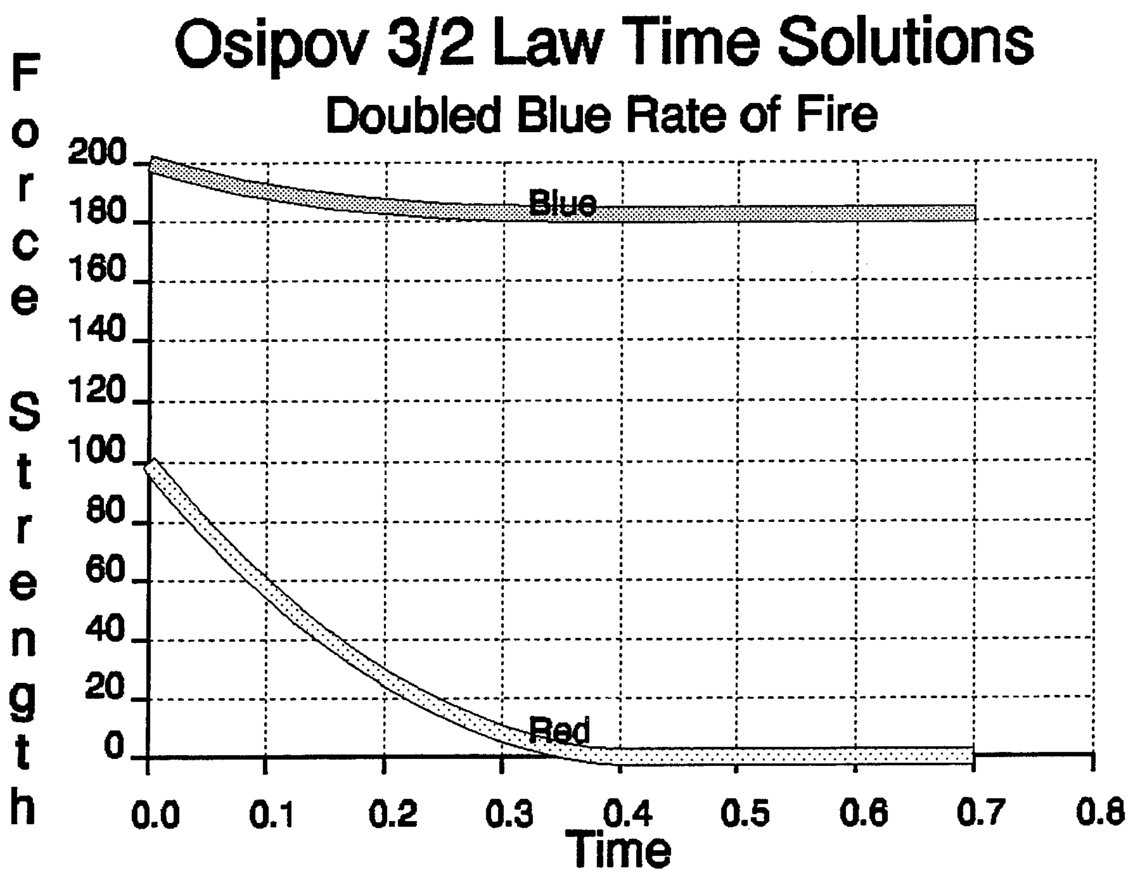


Figure (VII.F.2)

If we examine the effect of changing the attrition rates, we see that for a doubling of the Blue attrition rate (by doubling either r_f or p_k,) that the conclusion is indeed accelerated. This is shown in Figure (VIII.F.2). If we double both the rate of fire and the probability of kill, the conclusion is accelerated even more. This is shown in Figure (VIII.F.3). Note that in these succeeding cases, the losses for the Blue force decrease slightly and approximately linearly (a result of the initial choice of parameters,) with increase in attrition rate.

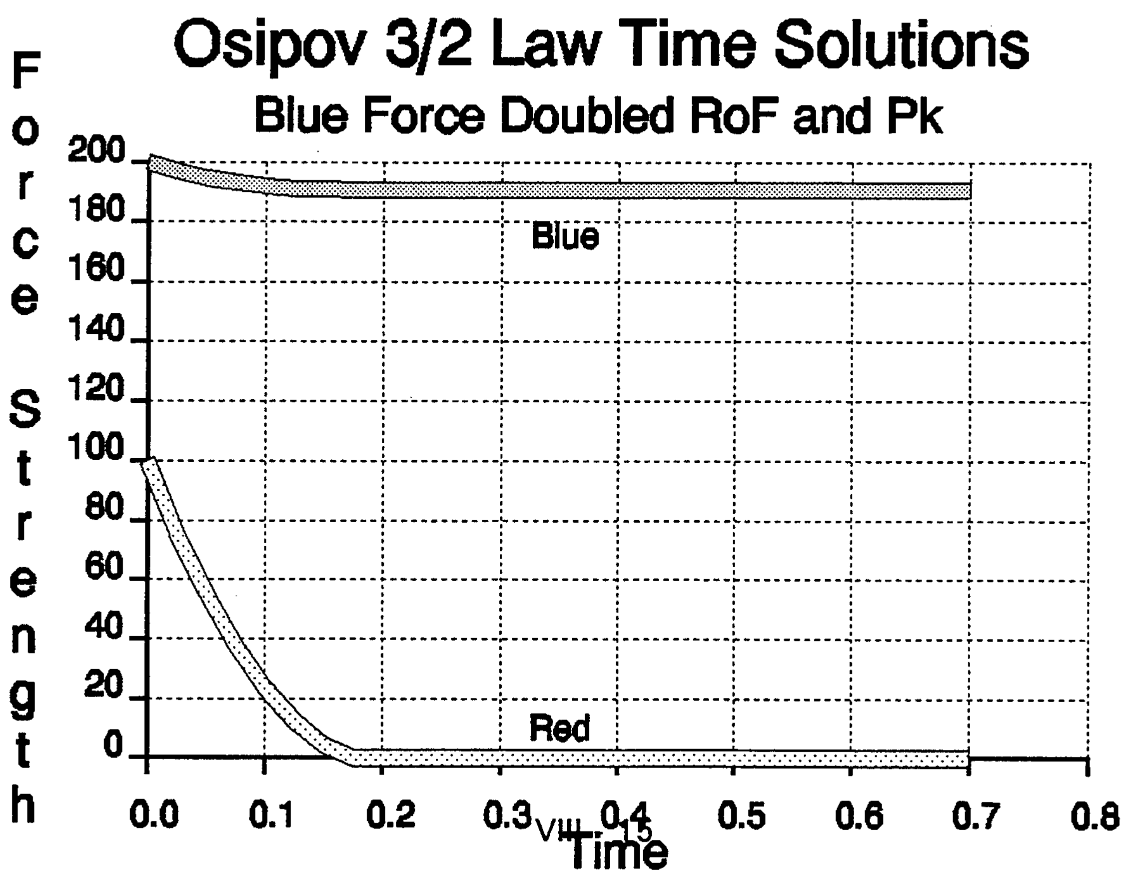


Figure (VII.F.3)

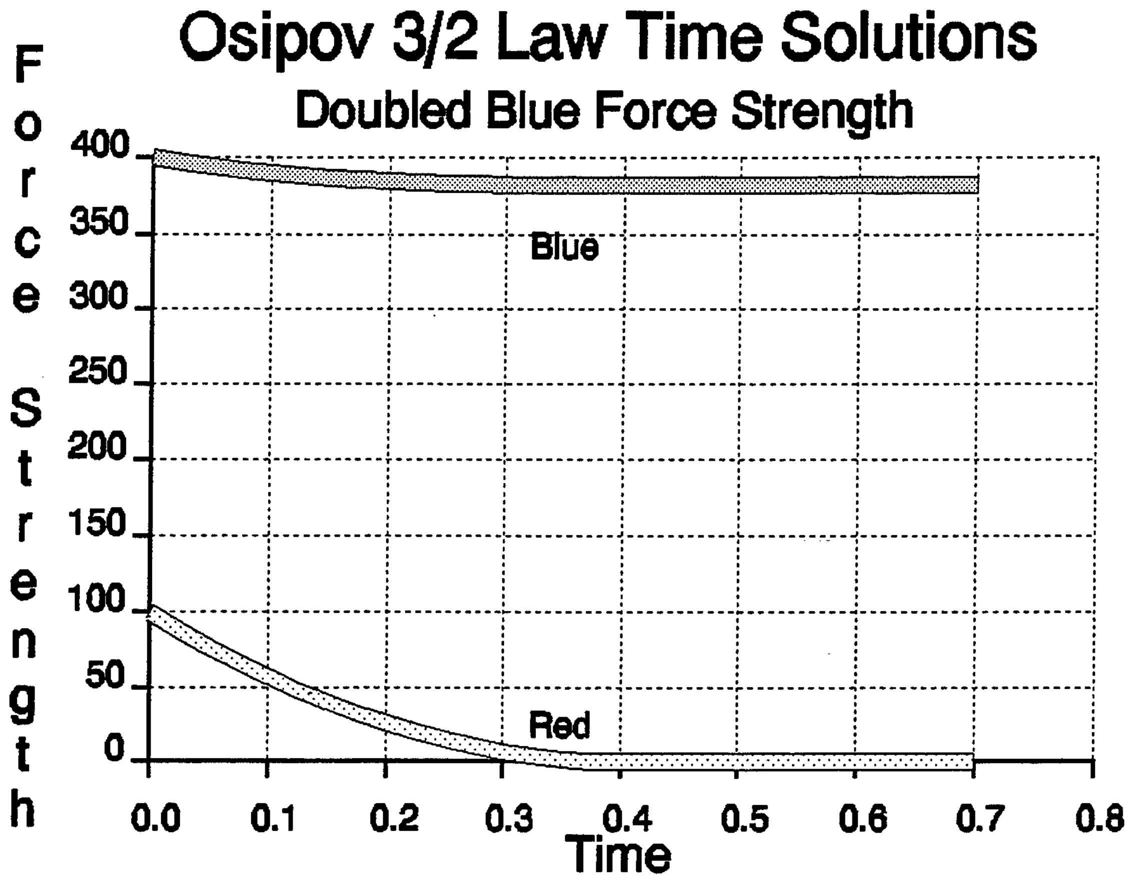


Figure (VII.F.4)

If we double (Figure (VIII.F.4)) and halve (Figure (VIII.F.5)) the initial Blue force strength, a much more dramatic change occurs in the form of the engagement. When we double the initial Blue force strength, the engagement concludes in approximately the same period of time as doubling the Blue attrition rate, and Blue losses are approximately the same. If we halve the Blue initial force strength, however, we find the engagement lasting approximately as long as the base case, but losses to the Blue force increase dramatically while losses to the Red force decrease. This is, of course, the draw case.

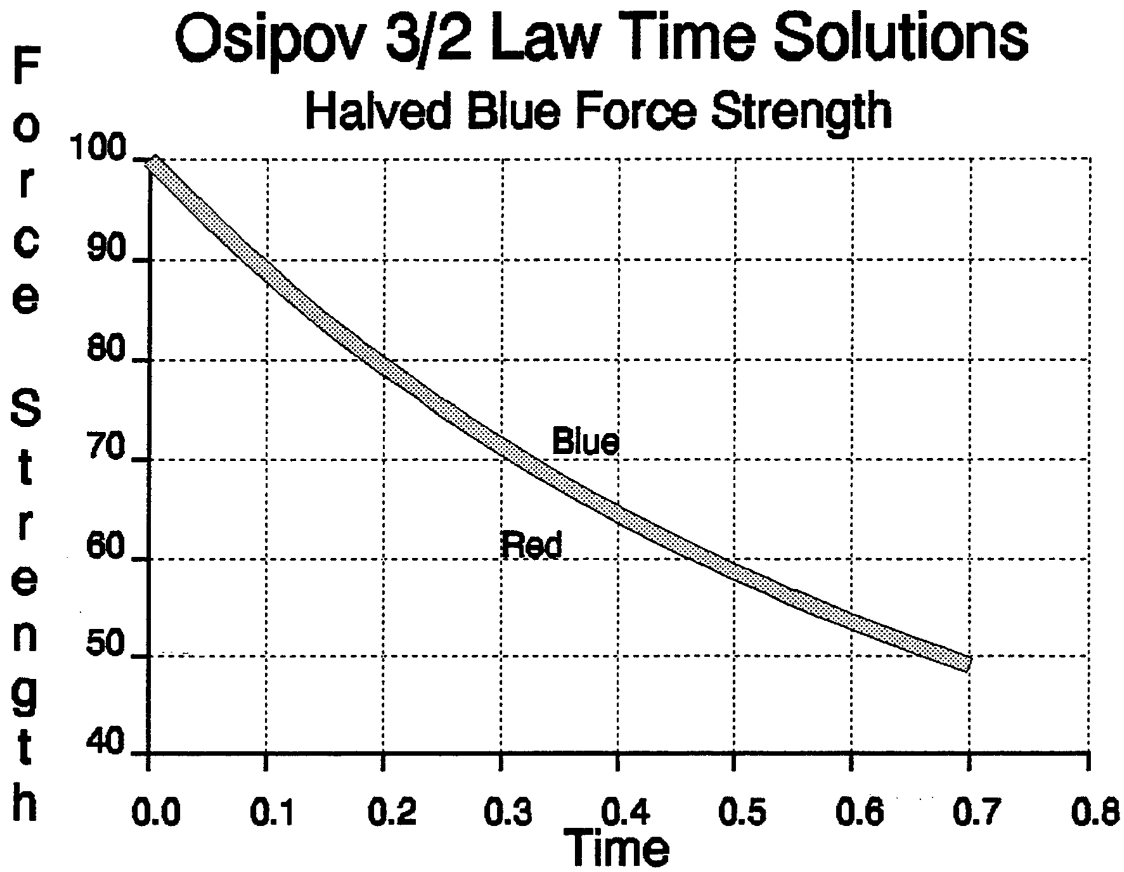


Figure (VII.F.5)

VIII.G. Comparison of the Square Law and Osipov's 3/2 Law

One of the interesting questions which comes up (in the next chapter for instance,) is a comparison of the Square and 3/2 Laws. The state solutions of the two laws are different, but the attrition rates are also. To examine this question we calculate the particular solutions for the square and 3/2 laws using the parameters of the preceeding section and the appropriate time solutions. This is shown in Figure (VIII.G.1). The student should keep in mind while examining this figure that while the initial force strengths, and the rates of fire and probabilities of kill are identical for the two calculations, the attrition rates are decidedly different. The square law attrition rates are $\simeq r_f\, p_k$, while the 3/2 law attrition rates are $\simeq r_f\, p_k / \sqrt{A_0}$. Thus the loss rate for the square law is $\simeq r_f\, p_k\, B$, while the loss rate for the 3/2 law is $\simeq r_f\, p_k \sqrt{A} / \sqrt{A_0}\, B$. Since during the initial stages of the engagement $A \simeq A_0$, we should expect little difference between the two solutions, while for longer time, we should expect that the 3/2 law solution will decrease slower than the square law solution. This is exactly what is shown in the figure.

The short time result is the more interesting of the two from a perspective of examining the historical data. Since most battles result in few casualties, it will be difficult to tell the difference between the square and 3/2 laws. Only by examining the correlation of the attrition rates with initial force strengths can we tell the difference between the two solutions.

VIII.H. Conclusion

In the Dawn of the Fourth Age: Aircraft in Warfare, Lanchester [,1916] makes the argument that ancient warfare was typified by the Linear Law. As a result of our investigations into the Ironman Analysis of General and 3/2 Osipovian Attrition, we are now in a position to hypothesize about the reasons for this.

If we postulate that the armies of the ancient period were drawn up in a line, we may make a comparison with the rank and file formation. From a mathematical standpoint, the line formation is a rank and file formation comprised of a single file. The depth of the formation is kept constant. From what we have described, this would lead us to expect the attrition to proceed according to Osipov's 3/2 Law.

This is not the case because the density of units along the front rank (the line) is not proportional to $\sqrt{A}$, but to A since the formation is always only one rank deep and the density of units in the line are either kept constant (and holes in the line are allowed to form,) or allowed to vary and thus the soldier - soldier ratio in contact changes. (The latter should shift to Square Law except that the holes allow the enemy to either pass through the line or redeploy and this is probably avoided.) Thus we may hypothesize that warfare in line gives rise to Linear Law attrition.

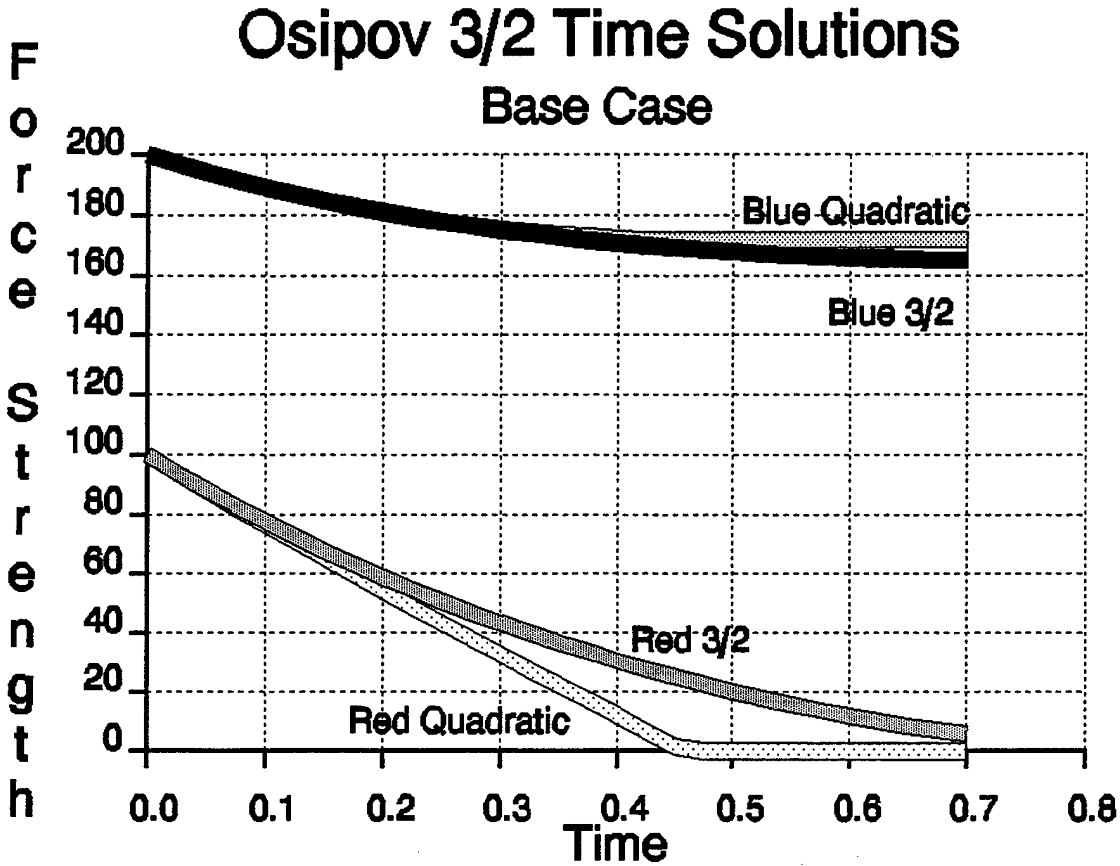


Figure (VIII.G.1)

IX. HISTORICAL PERSPECTIVES

IX.A. Introduction

One of the fundamental problems presented to the scientific study of attrition is the difficulty in acquiring data on the phenomena. First of all, experiments cannot be conducted - the very nature of war precludes deliberate scientific investigation on moral, ethical, philosophical and financial grounds. Numerous attempts have been made to simulate war in the form of training and test exercises but these have been flavored by the unrealities of referees, umpires, and rules, and the difficulty of gathering data in a meaningful manner. Recent efforts in analog combat simulation in this country, such as the National Training Center and the DARPA SIMNET Program offer some greater promise although, as Osipov would have possibly asserted, these simulations are primarily oriented to the teaching of military principles. Still even these simulations are flawed by the assumptions built into their underlying models.

This leaves, of course, historical information. Data on the battles of the past are difficult to find and different sources of these data may contradict each other. Even the barest minimum of data is available only for very few battles. As a minimum, four data are needed for each battle - the initial and final force strength for each of the two sides. A fifth datum, the duration of the battle, is hard to find in any precise terms. For extended (many day) battles, force strength as a function of time would be exceedingly valuable, but this type of data has been found for only a few battles.

Despite this lack, there have been several previous historical analyses. The first of these, recently rediscovered, was that of Osipov.[1] Other analyses include the Battle of Iwo Jima of World War II,[2] and the Inchon-Seoul campaign of the Korean-Chinese-American War (or Korean Police Action as it is often called in the United States).[3] We will review several such analyses in Section B.

The most elaborate of these appears to be that of Dupuy in the development of the Quantified Judgement Model (QJM),[4] but since most of that data base has not been submitted to open scientific scrutiny, the extent of the data base and the validity of the QJM remain unknown. (This last statement should not be construed as criticism of COL Trevor Dupuy and his co-workers. The QJM and its associated data base at the History Evaluation Research Organization (HERO) are financial ventures in the finest American tradition. They exist as profit makers and cannot be thereby expected to the part of the open scientific literature. Useful portions of the QJM and the database which have appeared in unclassified part may be used here as applicable and properly cited.)

As part of the analyses in this chapter, five primary sources have been used in compiling some short (unfortunately) databases of battle data. In addition to the data of the preceding section, Osipov's data,[1] and some data on World War I battles in Dupuy's Numbers, Predictions and War,[2] these compendia of warfare data were found: Eggenberger's **An Encyclopedia of Battles,**[5] describing 1,560 battles, Lafflin's **Brassey's Battles** covering some 7,000 battles,[6] and Livermore's **Numbers and Losses in the Civil War in America 1861-65.**[7] The data sets were gathered from each of these sources. There is some overlap of these data sets - that is, common battles. The force strengths enumerated by the authors/editors of these sources have been accepted without question. No resolution of conflicting numbers cited has been attempted. One special data base incorporating short battles ($\leq$ 1 day in duration) for which fairly exact durations are cited (in minutes or hours) has been compiled to permit some superficial analysis of the time dependence of the battles. Another special data base has been constructed for battles that proceeded to, or nearly to, conclusions. These battles are exceedingly rare, but do exist.

The author of this work makes no pretense of being an historian. The assumption has been made that the collectors of these data are historians of sufficient credentials that their compendia may be used to gain insight. The philosophical approach taken is essentially that originally espoused by Osipov, that in the aggregate, the historical data can give insight into the attrition phenomena. The student is therefore challenged to accept the contents of this chapter with care and deliberation; more so than much of the rest of this work which has a technical foundation. As we shall see in the next section, the previous Lanchester analyses of historical data have not strongly supported the Lanchester attrition theory. Indeed, it has only been recently that Lanchester theory has been conditionally accepted by historians.[8]

IX.B. Previous Historical Analyses. I

The open literature (I suspect now that the Soviet Union has dissolved, and their military analysis is more available to us, we shall find similar analyses on their parts) contains two historical Lanchestrian analyses of heroic form, analyses of individual battles. I characterize these analyses as heroic because of the difficulty of acquiring periodic (usually daily or hourly) strength figures for both sides. This is a far cry from the minimum four (initial and final) strengths that is our fundamental criterion for the data bases of the following sections.

Perhaps it is instructive that both of these battles/campaigns: Iwo Jima and Inchon-Seoul; are twentieth century so that record keeping has become a regular staff/historian activity. Perhaps the fact that one side involves Americans with their business-like approach to war transcends the geographic availability of the records. Nonetheless, the fact remains that these two analyses are cornerstones of quantitative military science, of the physics of war.

The Battle of Iwo Jima occurred late in the Second World War (Pacific Theater). The island of Iwo Jima, an 8 square mile, triangular shaped rock in the Bonine group was viewed as a threat to bomber operations from Saipan against the Japanese mainland. Its threat value was mirrored by its desirability as a forward air base[9].

The invasion, conducted from 19 February through 24 March of 1945 by the Fifth Marine Amphibious Corps (consisting of the 3rd, 4th, and 5th Division) MG Harry Schmidt commanding, and supported by the U.S. Fifth Fleet, was a classic 2-up, 1-back invasion which resulted in a successful occupation of the island by 11 March.

The Battle of Iwo Jima was a conclusive battle in the sense that the Japanese forces were completely (?) destroyed. (We will discuss conclusive battles in a subsequent chapter.) Although Japanese organized fighting was considered to have concluded on 16 March, by 11 March, their forces had been contained in two small coastal regions. In his analysis, Engel was able to obtain detailed data on American force strengths (arrivals and casualties) on a daily basis, and the duration of the different phases of the battle. He also knew the initial and final (presumed zero) strengths of the Japanese force strengths and that they were neither reinforced nor evacuated.

Table IX.B.1 American Force Strength Arrival Times

Day of Battle	U.S. Force Strength Increments
0	54000
1	0
2	6000
3	0
4	0
5	13000
6	0

The arrival times of American forces is given in Table IX.B.1

The battle lasted for 36 days and concluded with a total of 20,800 American casualties and 4,590 dead. The initial Japanese force strength was 21,500 (Depuy and Dupuy report 22,000).

Both Engel and Busse make the assumption that the Lanchester Quadratic attrition differential equations with reinforcements (on one side,) are applicable. No consideration is made of any attrition order other than two. Both forces are approximated as homogeneous. The relevant equations are thus

$$\frac{dA}{dt} = -\alpha J + a, \tag{IX.B-1}$$

and

$$\frac{dJ}{dt} = -\beta A. \tag{IX.B-2}$$

where: A(t), J(t) are the effective American and Japanese force strengths,
α, β are the attrition rates, and
a(t) is the American arrival rate.

The solution of these equations has already been described in Chapter VI.H. Because data are available only on a daily basis, a numerical approximation (finite difference) was necessary. Since the Japanese received no reinforcements, the relevant (approximate) force strength solutions may be written from equations (VI.H-20) and (VI. H-21) as

$$A(t+\Delta t) = A(t)\cosh(\gamma\Delta t) - \delta\, B(t)\sinh(\gamma\Delta t) + \frac{a(t)\sinh(\gamma\Delta t)}{\gamma}, \tag{IX.B-3}$$

and

$$J(t+\Delta t) = J(t)\cosh(\gamma\Delta t) - \frac{A(t)}{\delta}\sinh(\gamma\Delta t) - \frac{a(t)\,(\cosh(\gamma\Delta t) - 1)}{\alpha}. \tag{IX.B-4}$$

The parameters γ and δ are (again) defined in the usual manner.

The attrition rates for the entirety of the battle and the campaign were estimated by integrating Equations (IX.B-1) and (IX.B-2) numerically over the entire duration as

$$A(\tau) - A_0 \approx -\alpha \sum_0^\tau B(t') + \sum_0^\tau a(t'), \quad \text{(IX.B-5)}$$

and

$$B(\tau) - B_0 \approx -\beta \sum_0^\tau A(t'), \quad \text{(IX.B-6)}$$

where τ is the duration of the battle/campaign, and all sums run from t' = 0 to t' = τ. The attrition rates are thus,

$$\alpha \approx \frac{A_0 - A(\tau) + \sum_0^\tau a(t')}{\sum_0^\tau B(t')}, \quad \text{(IX.B-7)}$$

and

$$\beta \approx \frac{B_0 - B(\tau)}{\sum_0^\tau A(t')}. \quad \text{(IX.B-8)}$$

The attrition rates are thus averaged over the entire battle - it assumes the attrition pace of the battle is constant. While the student may view this as a strong and possible enormous assumption, reflection should renewal that this is almost a forced assumption. Given the nature of the data available and the conclusions we have made thus far in Lanchester attrition theory, the assumption of constant attrition rate is logical and natural.

Since the data do not include daily Japanese force strengths, J(t) can only be calculated by recourse to the same attrition differential equations. There are no actual daily Japanese force strengths for comparison. There are actual daily American force strengths available (calculable from daily arrival and casualty data for comparison although Engel does not explicitly include these in his article.)

Since Engel already had the actual daily American force strengths, he could examine the estimated daily American force strengths with the actual numbers even though he did not have the same detailed data on the Japanese. This comparison constitutes a strong test of Lanchester attrition theory with its fundamental assumption of constant attrition rate.

Needless to say, the comparison was made, and Engel found very close agreement with the actual data. Admittedly, there are daily fluctuations - differences between the actual and numerically estimated force strengths, but these differences were small percentages. The key point is not the attrition rates were assumed constant, but that the battle was simulated in a cumulative manner dictated by Lanchester attrition theory. This cummulation process tends to accumulate all the errors, inaccuracies, and differences. Thus, the simulation of the n^{th} day of combat carries with it all of the errors and differences that the model has generated on all of the previous n-1 days of combat.

Admittedly, we would expect some of these errors and differences to cancel, but the degree of consistent agreement that Engel found is a telling demonstration of the validity of Lanchester attrition theory. Clearly, this analysis is a strong argument that the Battle of Iwo Jima, to the degree that we have descriptive actual data, is described by Lanchester attrition theory.

As a final note, Engel notes that there are other models that one could assume to analyze these data, and that these models could also have good agreement with the data, the model that he presents here is the simplest of these. If we place our trust in Occam's Razar, then this simplest of models is the valid one.

Before we proceed to the next analysis, it seems worthwhile to comment on the assumption of constant attrition rates. Recall that Engel did not have detailed data on the daily Japanese force strengths, only their initial and final force strengths. If he had been able to get these data, then he could have used daily (rather than battle) averaged force strengths and have improved the agreement between actual and estimated force strengths. I will contend that while this may have reduced the difference between actual data estimates, it would only have clouded his actual contribution. Engel used the framework of Lanchester attrition theory to perform this analysis. The agreement of calculation with actual data constitutes reasonable demonstration of the applicability of the model.

Obviously, the data analysis weakness of Engel's analysis is the lack of detailed Japanese daily force strengths. The second analysis that we describe here, of the Inchon-Seoul Campaign reported by Busse, set out to address that very deficiency by using estimated enemy daily casualties (force strength) derived from intelligence.

The Inchon-Seoul campaign, 15-26 September 1950, began with the amphibious launching of the American X Corps (1st Marine Division, 7th Infantry Division), MG Edward L. Almond commanding, at Inchon, Korea, on the Yellow Sea, and culminated with liberation of Seoul. This daring campaign, began by an amphibious landing like the Iwo Jima battle, enabled General Douglas MacArthur to regain the initiative, destroy the North Korean Army, and advance north to the Yellow River until the Chinese counteroffensive begun on 25 November.

The basic data for Busse's analysis are given in Table (IX.B2).
Table IX.B.2. Inchon-Seoul Force Strengths and Reinforcements

Day	Marine Force Strength	North Korean Reinforcements	North Korean Force Strength
0	25040	0	22150
1	24844	0	21350
2	24818	0	20500
3	24742	3000	22750
4	24640	500	22600
5	24568	230	22100
6	24421	6500	27675
7	24190	0	25975
8	24025	0	24375
9	23882	2000	25305
10	23593	0	24290
11	23317	0	22390
12	23114	3500	24640
13	22925	0	23250
14	22882	0	22710
15	22813	0	22100
16	22752	2000	23465
17	22733	5000	28265
18	22636	0	27835
19	22598	0	26930

The relevant attrition differential equations are

$$\frac{dK}{dt} = -\beta A + k(t), \tag{IX.B-9}$$

and

$$\frac{dA}{dt} = -\alpha K,$$

where: A(t), K(t) are the American, (North) Korean force strength,
α,β are (again) the attrition rates, and
k(t) is the (North) Korean reinforcement rate.

Again in this case, only one side receives reinforcements, and since the data are daily in nature, we may write the numerical approximations for the force strength as

$$\begin{aligned} A(t+\Delta t) &= A(t)\cosh(\gamma\Delta t) - \delta K(t)\sinh(\gamma\Delta t) \\ &\quad - \frac{k(t)}{\beta}(\cosh(\gamma\Delta t) - 1), \end{aligned} \tag{IX.B-11}$$

and

$$\begin{aligned} K(t+\Delta t) &= K(t)\cosh(\gamma\Delta t) - \frac{A(t)}{\delta}\sinh(\gamma\Delta t) \\ &\quad + \frac{k(t)\sinh(\gamma\Delta t)}{\gamma}, \end{aligned} \tag{IX.B-12}$$

from Equations (VI. H-20) and (VI. H-21). The parameters γ and δ are (again) defined in the usual manner.

Busse, like Engel, assumes a campaign averaged attrition rates. Since he has actual daily force strengths for both sides, Busse may calculate both of his attrition rates directly from numeric integration of the attrition differential equations.

Since the values of α and β are small compared to Δt, Busse approximates the cosh and sinh with one term Taylor series so that Equations (X.B-11) and (X.B-12) become

$$A(t+\Delta t) \approx A(t) - \alpha K(t)\Delta t, \tag{IX.B-13}$$

and

$$K(t+\Delta t) \approx K(t) - \beta A(t)\Delta t + k(t)\Delta t. \tag{IX.B-14}$$

The results of the calculations are shown in Figures (IX.B.1) - (IX.B.2). We note good agreement during the early part of the campaign, but decreasing agreement as the campaign progresses. This is to be expected due to the accumulation of errors in the method of numerical integration, and the approximation of constant attrition rates.

Busse did not consider his analysis to be a general validation of Lanchester theory because he had to use a finite difference form of the solution. This is a strong point, but it is not compromising. The finite difference form is derived from the Lanchester Theory and thereby partakes of its assumptions and limitations. No mathematical theory may be applied to real data without approximation. Thus, the results of both Engel and Busse, while not conclusive, must be considered to support the applicability of Lanchester attrition theory.

IX.C. Previous Historical Analyses. II

As we earlier stated, there have been several other analyses of historical data in terms of Lanchester theory. If, for the moment, we exclude Osipov's analysis, and the analysis of the preceding section, we find two analyses, both cited in Dupuy's **Numbers, Predictions, and Wars**. These analyses are due to Dr. Daniel A. Williard, and Janice B. Fain. These analyses found that the attrition order of historical battles was approximately 2.5. In terms of the attrition differential equations, these analyses indicate a form of

$$\frac{dA}{dt} = -\frac{\alpha}{\sqrt{A}}\,B\,. \qquad \text{(IX.C-1)}$$

which would lead us to believe that the rate of loss is inversely proportional to the square root of the force strength - the stronger a force the fewer casualties it takes. This result is used by Dupuy to argue the invalidity of Lanchester's theory. We note in passing that Equation (IX.C-1) is an Osipov type of attrition differential equation and that we can and do have an explicit time solution of this type of (paired) equation. We further note that we might subject the data to an Osipov type analysis for an attrition order of 5/2 and find good agreement - we will defer such considerations for the moment.

The data in Williard's paper and the first Fain paper are not available for this work because of their limited nature (we restrict ourselves to open literature sources). The second Fain paper is, however, openly available and we may examine it to gain some insight into that analysis. The data used in this study were extracted from Bodart which was not available to this author, and the report does not exhaustively catalog the battles in Fain's data base (approximately 1100 battles). For our calculations here, we substitute other data bases drawn from other sources. We may only assume that they are similar.

Marine Force Strength

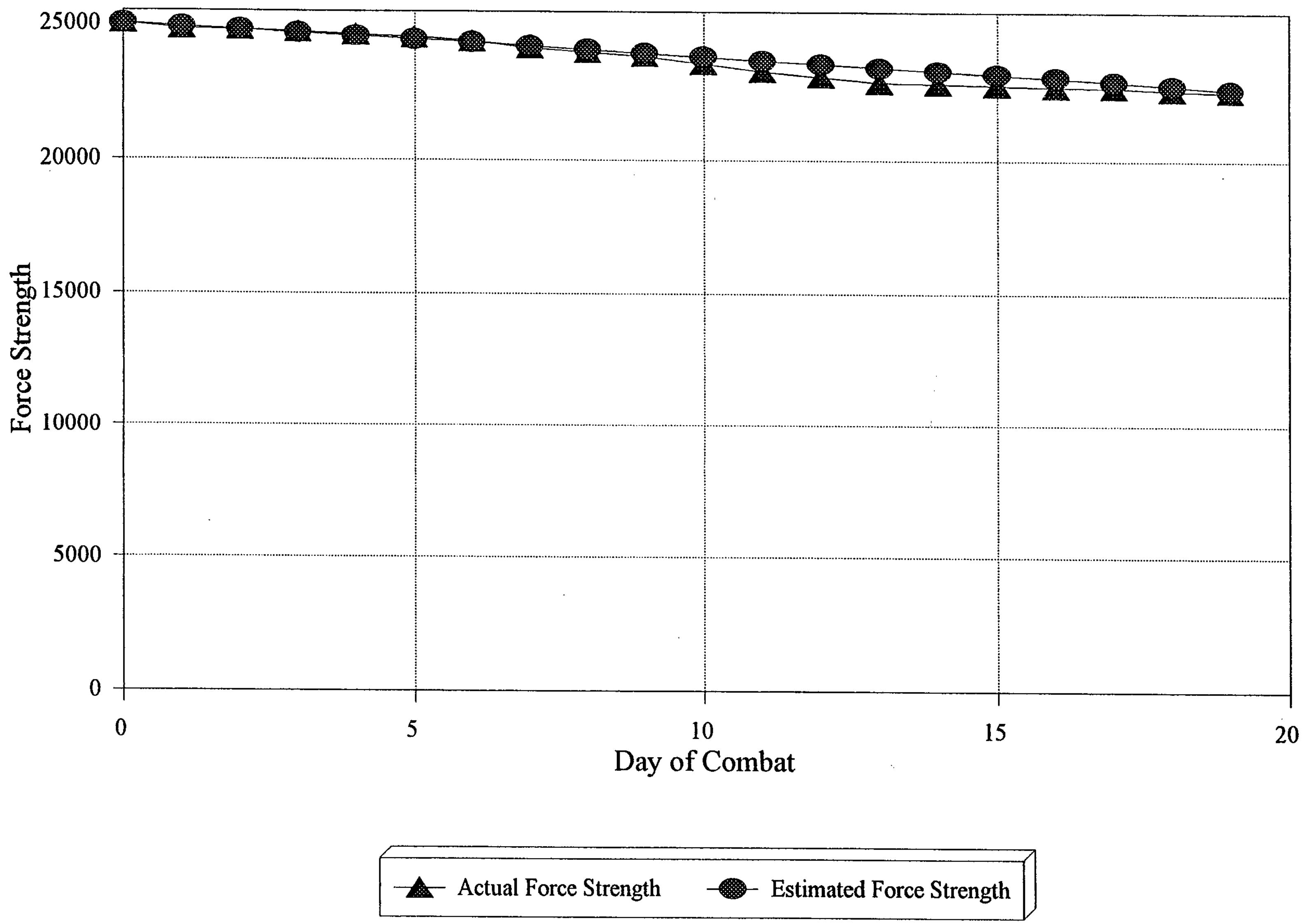


Figure (X.B.1)

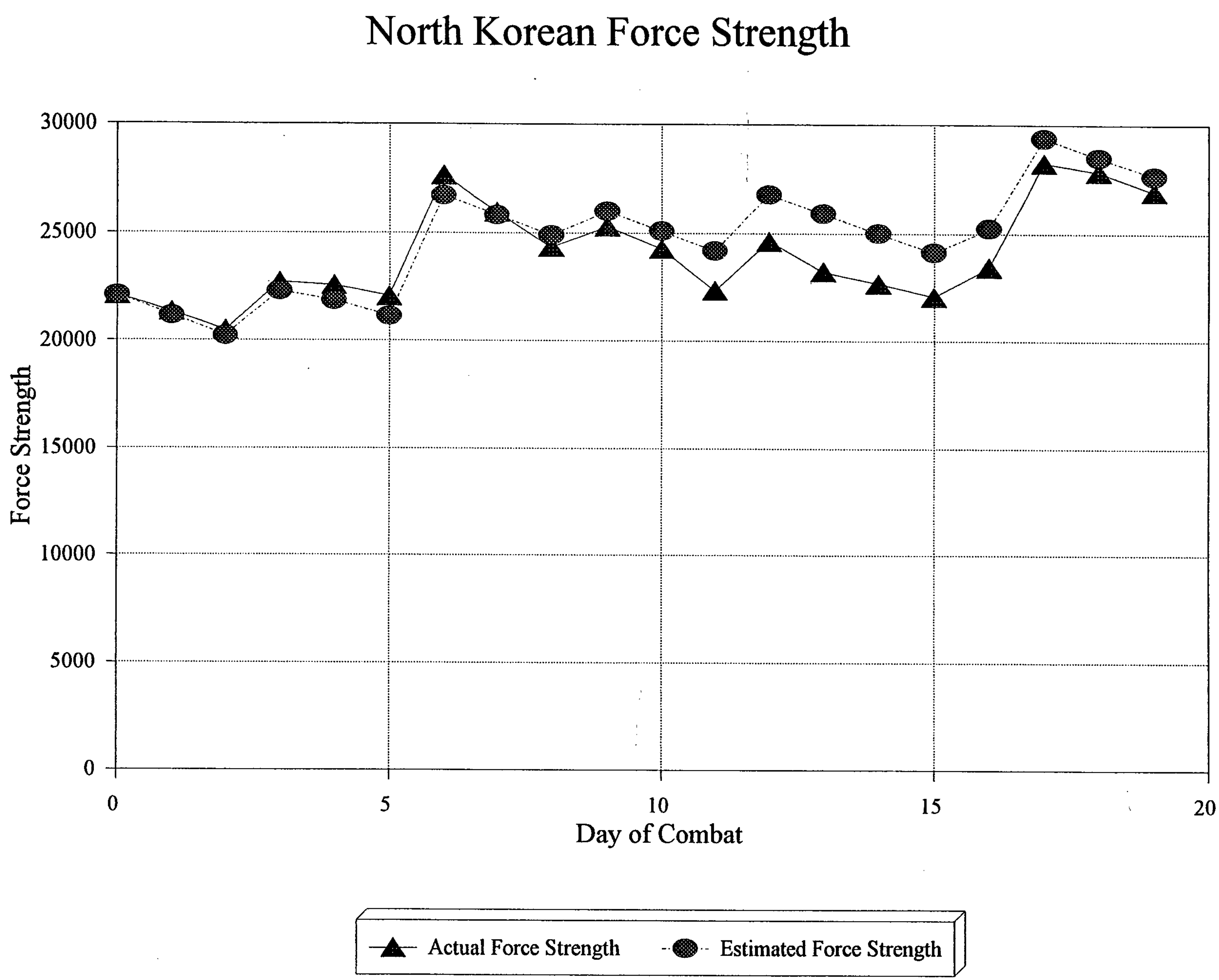

North Korean Force Strength
30000
25000
20000
15000
10000
5000
0
Force Strength
0
5
10
15
20
Day of Combat
Actual Force Strength
Estimated Force Strength

Figure (IX.B.2)

The analyses of Williard and Fain were based on the idea of performing linear regression of what is essentially the state solution. (We translate this method into our own notation here for consistency sake). The Lanchester-Osipov state solution is

$$\beta\ (A_0^n - A^n) = \alpha\ (B_0^n - B^n). \tag{IX.C-2}$$

We may rewrite this as

$$A_0^n \left(1 - \frac{A^n}{A_0^n} \right) = \frac{\alpha}{\beta} B_0^n \left(1 - \frac{B^n}{B_0^n} \right). \tag{IX.C-3}$$

As Osipov noted, and reiterated by Fain, most battles end when losses are still small (20-30%). It is therefore convenient to rewrite the final numbers using the loss functions introduced by Osipov. In those terms

$$\begin{aligned} A &= A_0 - a \\ B &= B_0 - b, \end{aligned} \tag{IX.C-4}$$

where a,b are the losses of the red, blue forces.

We may write Equation (IX.C-3) as

$$A_0^n \left[1 - \left(1 - \frac{a}{A_0} \right)^n \right] = \frac{\alpha}{\beta} B_0^n \left[1 - \left(1 - \frac{b}{B_0} \right)^n \right], \tag{IX.C-5}$$

and since a/A_0, b/B_0 are small compared to 1, we may write

$$A_0^{n-1}\ a = \frac{\alpha}{\beta} B_0^{n-1}\ b, \tag{IX.C-6}$$

after cancelling a common factor of n. If we take the logarithm of the equation, we get

$$(n-1)\ln(A_0) + \ln(a) = \ln\left(\frac{\alpha}{\beta}\right) + (n-1)\ln(B_0) + \ln(b), \tag{IX.C-7}$$

which we may rearrange as
which we immediately see has the form of a straight line

$$\ln\left(\frac{a}{b}\right) = \ln\left(\frac{\alpha}{\beta}\right) + (n-1)\ln\left(\frac{A_0}{B_0}\right), \qquad \text{(IX.C-8)}$$

$$Y = c + dx,$$

where c is the intercept and d is the slope of the line. If we curve fit an historic data base, where Y = ln(a/b), and x = $\ln(B_0/A_0)$ we can expect to find the fitted slope to be related to the attrition order.

From the sources available to us a series of data bases were constructed. Even taken together, these data bases contain fewer total battles than does Fain's data base. These data bases consist of the following:[a]

A data base which includes those battles with the requisite five data per battle, with battle duration expressed in days. These battles (108 in number) were drawn from Laffin and Eggenberger. Most of these battles were one day in duration - few are of more than 6 days in duration. These data are presented in Table (D.1) and will be referred to as Nominal Length Battles.

The second data base was drawn from Livermore. Duration was calculated in days based on the inclusive dates of the battle. This data base consists of 49 battles, and is presented in Table (D.2), and will be referred to as Civil War Battles

The next data base is drawn directly from Osipov. It does not include duration. This data base is given in Table (D.3), and will be referred to as Osipov's Battles.

The fourth data base is somewhat arbitrarily formed of short battles, most less than one day in duration. An upper and lower bound on the duration of the battles (in hours) was placed in an attempt to facilitate the calculation of time dependence. These data are given in Table (D.4), and will be referred to as Short Battles.

The fifth data base, consisting of World War I battles, is presented in Table (D.5). There are all relatively long duration battles.

A sixth data base was also developed for battles which were fought to, or near to, a conclusion. These battles will be discussed in a later chapter.

Each of these data bases was subjected to a correlation analysis and to a linear

---

[a] These databases are presented explicitly in Appendix D. I apologize for the inconvenience to the student, but since we shall refer to these data in later chapters as well, and I do not want to continually reproduce the same data, I have chosen to locate them at the end of the book.

regression analysis ala Equation (IX.C-5). The results of that analysis are presented in Table (X.C.1)

Table IX.C.1 Data Set Analysis Using Willard's Formula

DATA SET	ATTRITION ORDER	ERROR	CORRELATION	# BATTLES
Nominal Length Battles	0.34	0.13	-0.44	108
Civil War Battles	-0.33	0.26	-0.53	49
Osipov's Battles	0.80	0.61	-0.05	45
Short Battles	0.58	0.20	-0.24	72
WWI Battles	0.02	0.36	-0.65	12

From this table, we see that the attrition order varies enormously from one data set to another. They evidently do not present an attrition order of 2.5. Lacking the data sets of Williard and Fain, we cannot speculate further on the discrepancies between these results and those of these two workers. We can, however, examine the validity of Equation (IX.C-8), which we shall do in the next chapter.

IX.D Further Analysis of the Data Sets.

In the process of performing the correlation analysis of the data sets, a curious feature was noted. The initial and final force strengths were found to be highly correlated. This is shown in Table (IX.D.1)

Table IX.D.1 Initial-Final Force Strength Correlations

DATA SET	A_0-A Correlation	B_0-B Correlation
Nominal Length Battles	0.98	0.98
Civil War Battles	0.99	0.99
Osipov's Battles	0.98	0.98
Short Battles	0.99	0.99
WWI Battles	0.98	0.98

These correlations are entirely too strong to be ignored. To illustrate this, the final and initial force strength are plotted in the following figures shown collectively in Table (IX.D.2).

Table IX.D.2 Correlation Plots of the Data Sets

DATA SET	FIGURES	FIGURES
Nominal Length Battles	IX.D.1 (Blue Force)	IX.D.2 (Red Force)
Civil War Battles	IX.D.3 (Union)	IX.D.4 (Confederate)
Osipov's Battles	IX.D.5 (Stronger Force)	IX.D.6 (Weaker Force)
Short Battles	IX.D.7 (Attacker)	IX.D.8 (Defender)
WWI Battles	IX.D.9 (Attacker)	IX.D.10 (Defender)

Examination of these figures reveals a high degree of linearity of the data. There is some scatter, part of which may be due to the uncertainties in the basic data. This degree of uncertainty, however, is not sufficient to negate the obvious conclusion that there is some (fairly simple) linear relationship between initial and final force strengths.

It is also possible to postulate from this data that there is more than one linear relationship between the force strengths. In Figure (IX.D.1), a decided slope change may be seen for initial force strengths >150,000 as compared to initial force strengths less than this value. A similar, but less well defined situation seems to exist for Figure (IX.D.2) (It is moot to draw too strong a conclusion for these nominal length battles as we draw no distinction between winner/looser or attacker/defender.)

As we proceed to examine the other figures, we observe no slope change for the Civil War Battles (Figures IX.D.3 and IX.D.4), but none of these battles have initial force strength > 120,000. If we examine Osipov's Battles (recalling now that force strengths are given in thousands.) no clear slope change occurs for the Stronger Force until ~ 250,000 (Figures IX.D.5 and IX.D.6). For the Weaker Force, there is a weak slope change for ~ 175,000. This contrasts with Osipov's conjuncture that attrition order changes at ~ 175,000 force strength. For the Short Battles, (Figures IX.D.7 and IX.D.8), there are weak slope changes for force strength ~ 100,000. The World War I data are shown in Figures IX.D.9 and IX.D.10.

From these plots, we may conclude that there is a strong linear relationship between initial and final force strengths and that the slope of this relationship may change as the initial force strength increases (changing at 100,000 - 150,000). The exact quantification of these relationships may be calculated using linear regression,

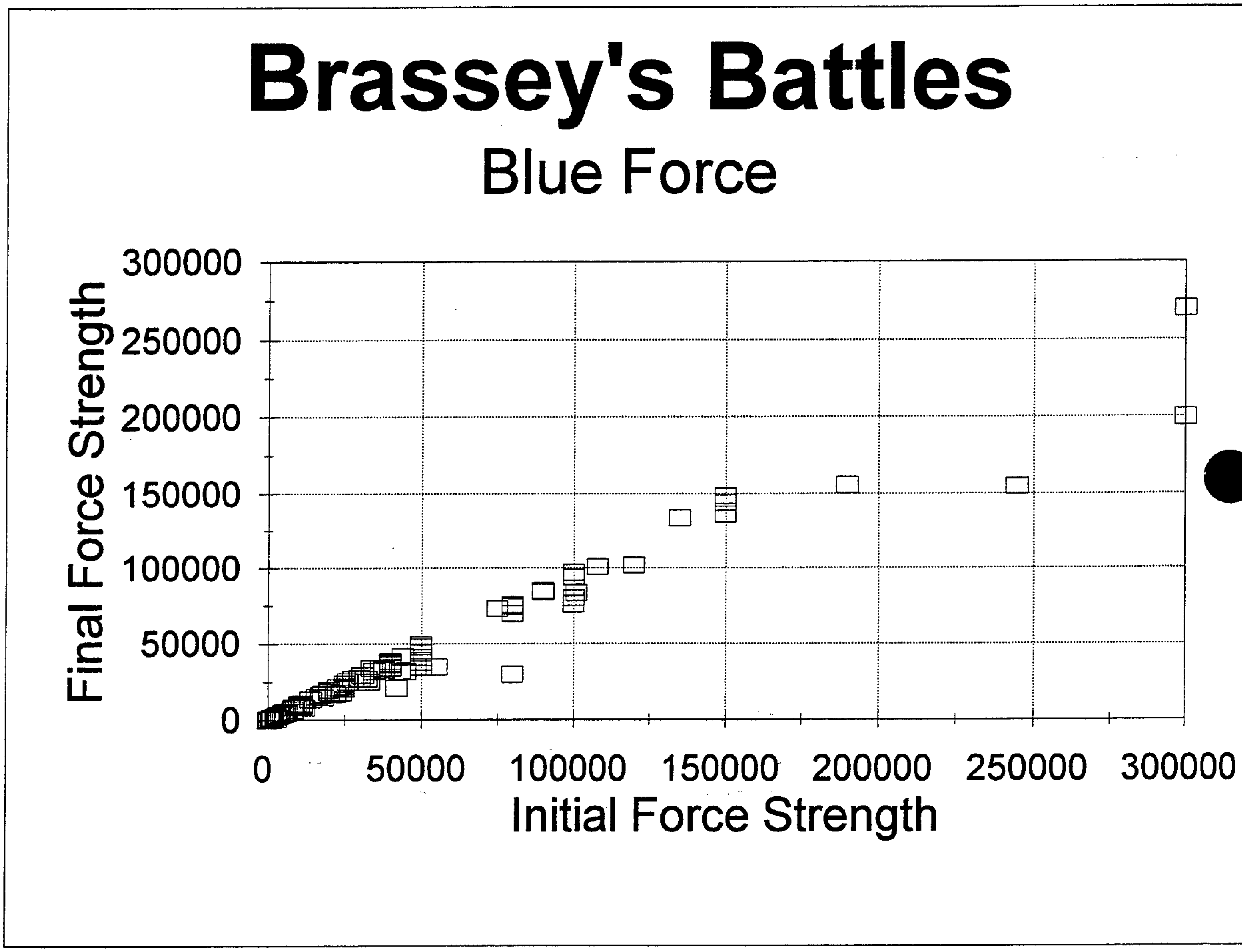

Brassey's Battles
Blue Force
Final Force Strength
300000
250000
200000
150000
100000
50000
0
0
50000
100000
150000
200000
250000
300000
Initial Force Strength

Figure IX.D.1

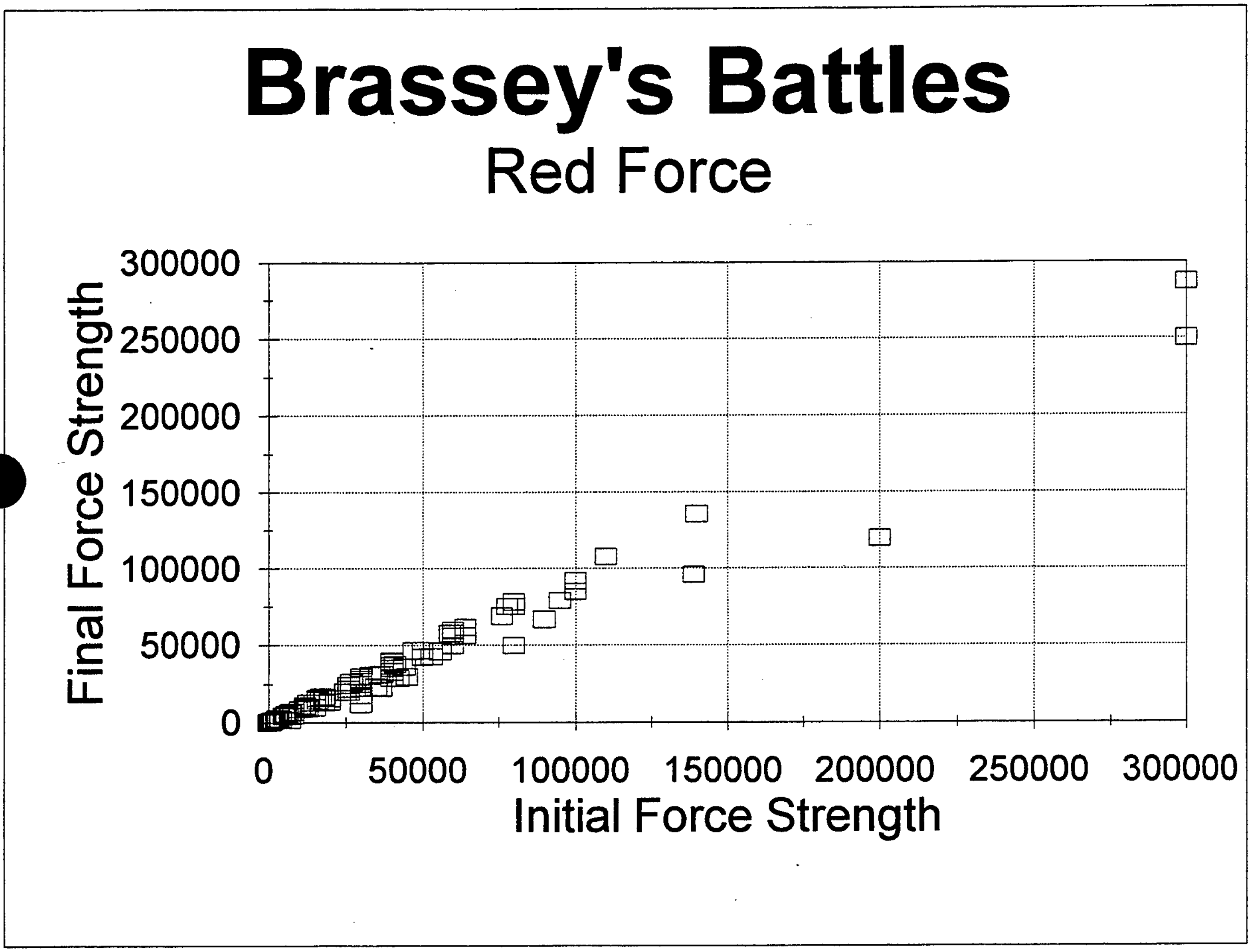

Brassey's Battles
Red Force
300000
250000
200000
150000
100000
50000
0
0
50000
100000
150000
200000
250000
300000
Final Force Strength
Initial Force Strength

Figure IX.D.2

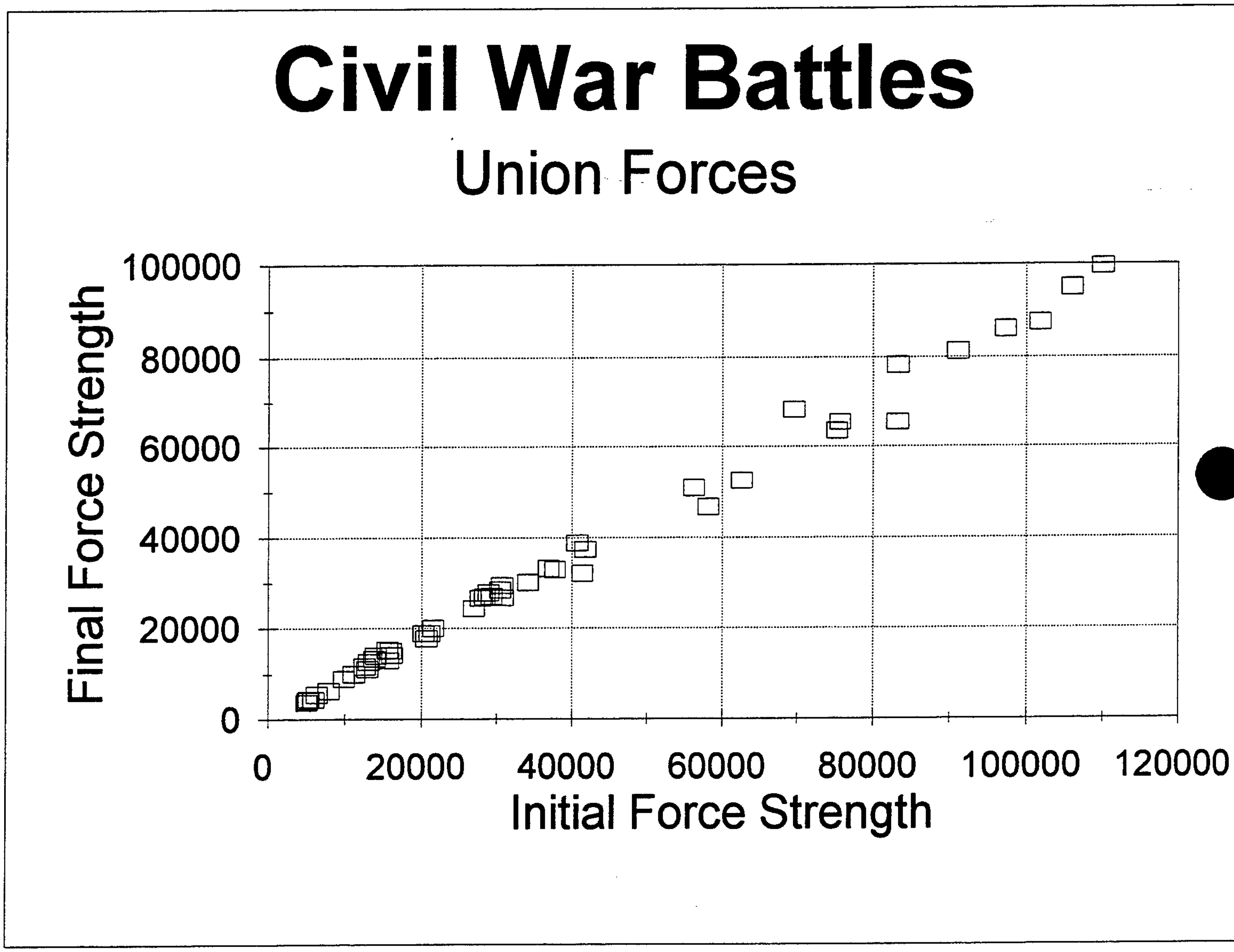

Civil War Battles
Union Forces
Final Force Strength
100000
80000
60000
40000
20000
0
0
20000
40000
60000
80000
100000
120000
Initial Force Strength

Figure IX.D.3

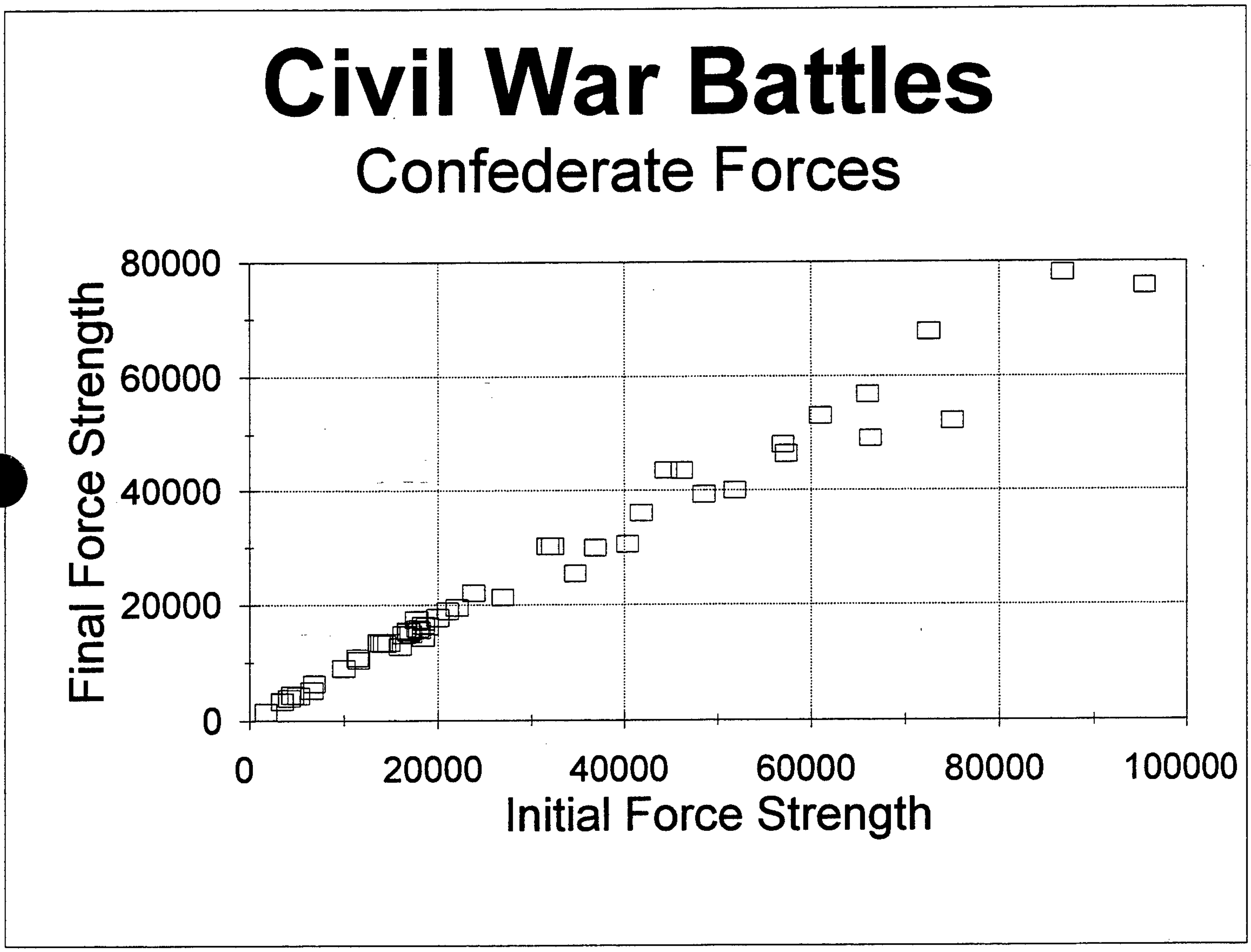

Civil War Battles
Confederate Forces
Final Force Strength
Initial Force Strength
0
20000
40000
60000
80000
100000

Figure IX.D.4

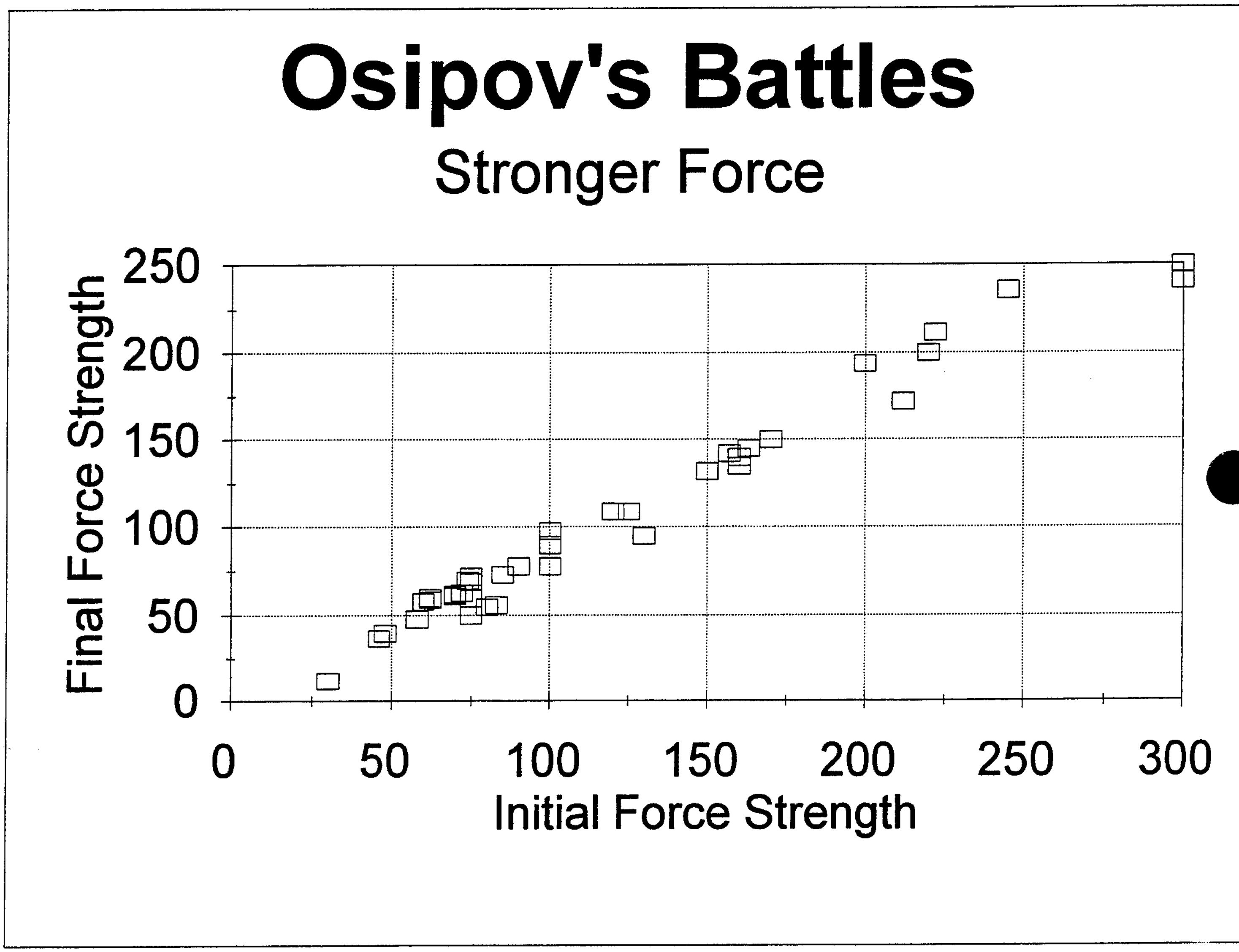
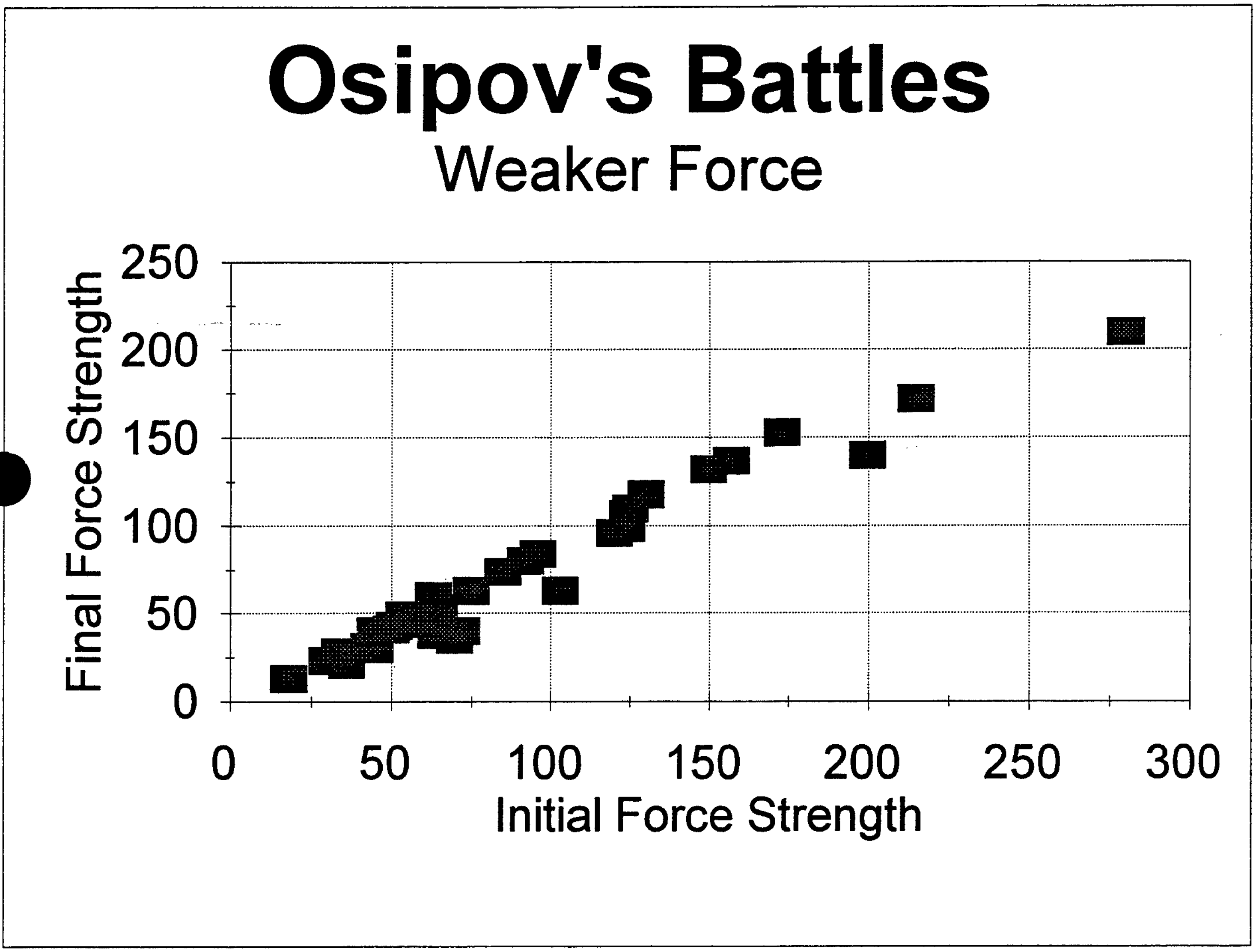

Osipov's Battles
Stronger Force
Final Force Strength
Initial Force Strength
0
50
100
150
200
250
300

Figure IX.D.5



Osipov's Battles
Weaker Force
Final Force Strength
Initial Force Strength
0
50
100
150
200
250
300

Figure IX.D.6

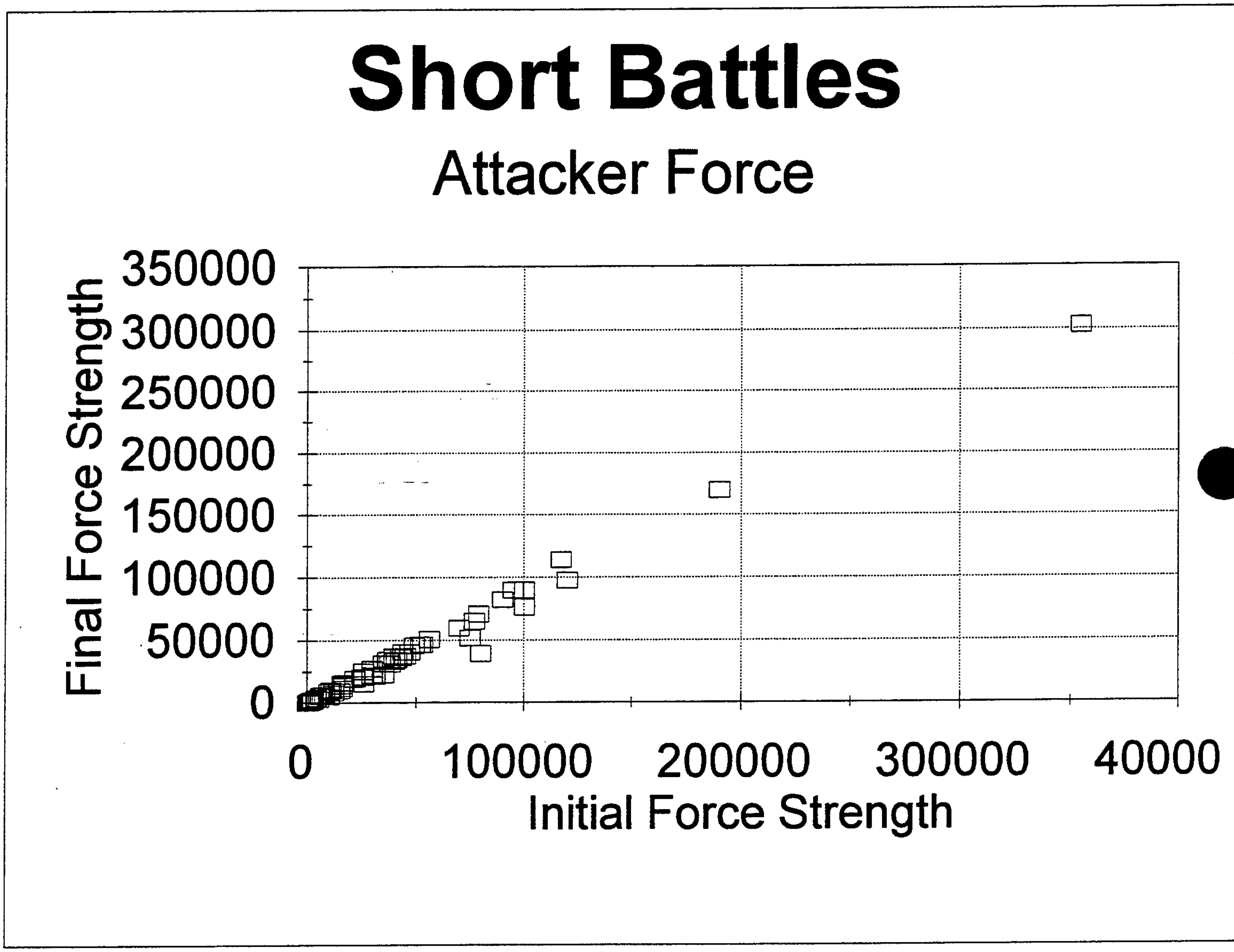

Short Battles
Attacker Force
350000
300000
250000
200000
150000
100000
50000
0
Final Force Strength
0
100000
200000
300000
40000
Initial Force Strength

Figure IX.D.7

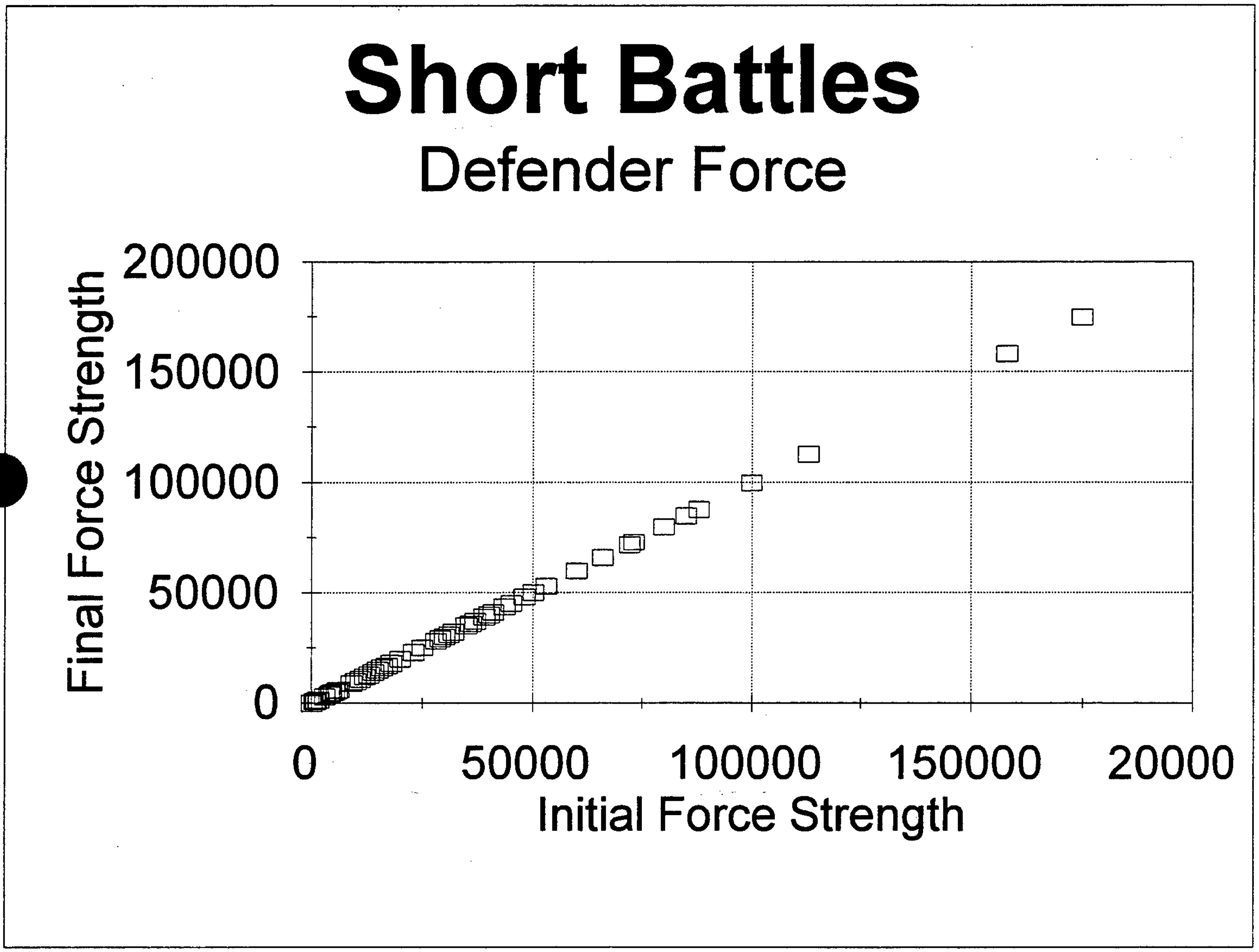

Short Battles
Defender Force
Final Force Strength
200000
150000
100000
50000
0
0
50000
100000
150000
20000
Initial Force Strength

Figure IX.D.8

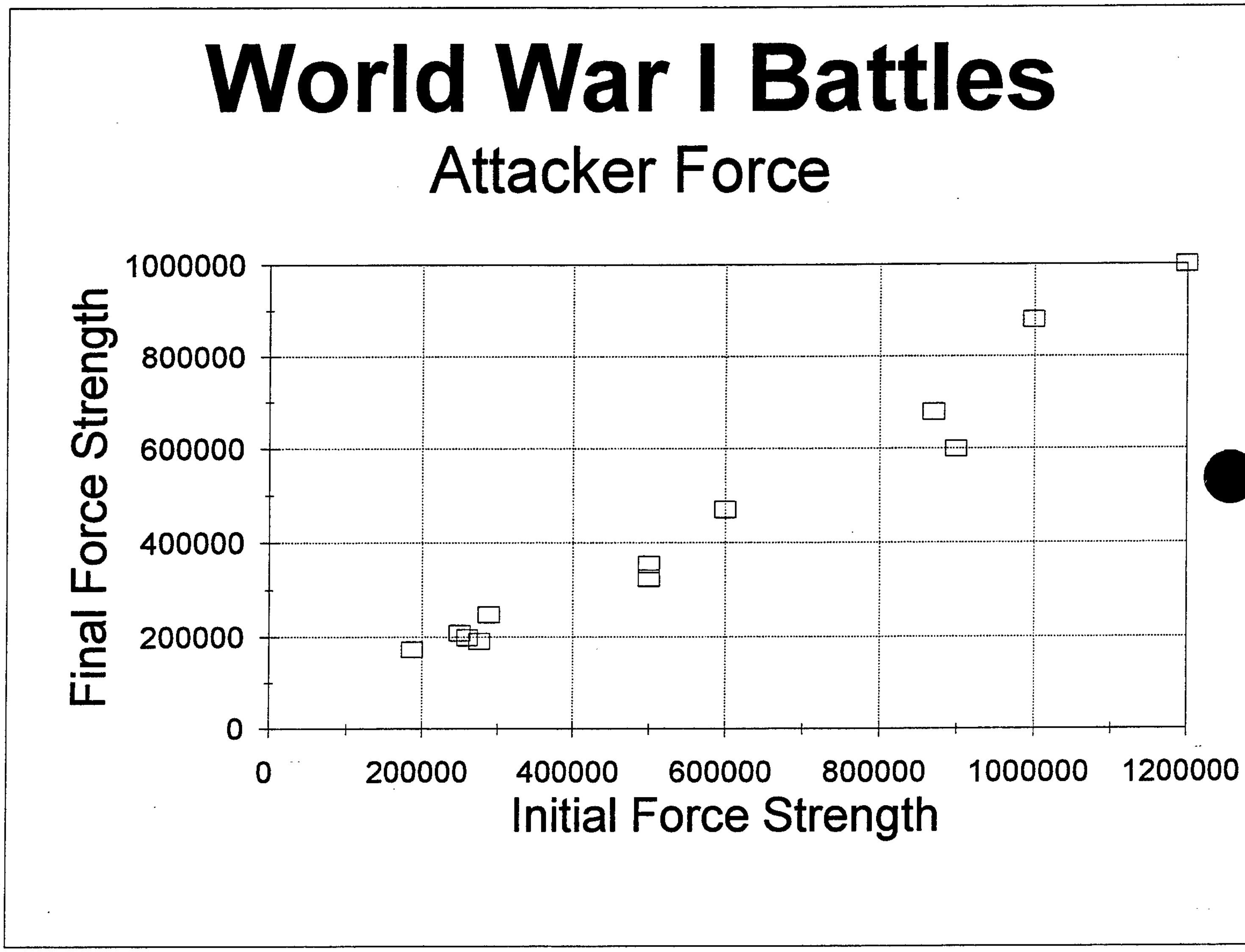

World War I Battles
Attacker Force
Final Force Strength
Initial Force Strength
0
200000
400000
600000
800000
1000000
1200000

Figure IX.D.9

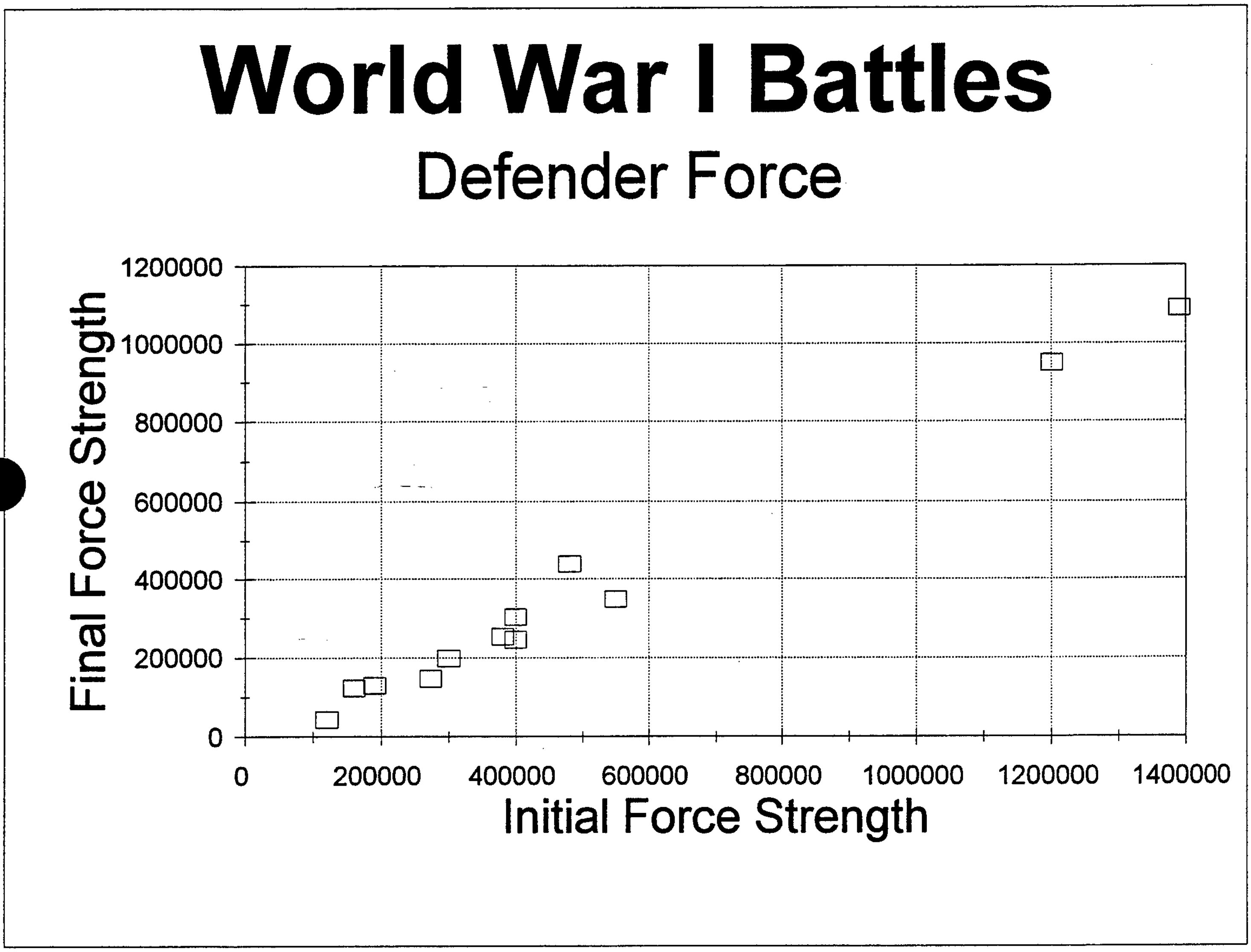


Figure IX.D.10

but before we perform such an analysis, a short side trip is needed.

References

1. Osipov, M., "The Influence of the Numerical Strength of Engaged Forces on their Casualties", Voennie Shornik (Military Collection), 3-7, 1915.

2. Engel, J. H., "A Verification of Lanchester's Law", Journal of Operations Research, 2 163-1[illegible], 1954.

3. Busse, James J., "An Attempt to Verify Lanchester's Equations", Paper 9-1, pp. 587-597, in Benjamin Avi-Itzhak, ed., **Developments in Operations Research**, Vol. 2, Gordon and Breach, New York, 1969.

4. Dupuy, Trevor N., COL (USA Ret.), **Numbers, Predictions, and War**, Bobbs-Merrill, New York, 1979.

5. Eggenberger, David, **An Encyclopedia of Battles**, Dover Publications, Inc., New York, 1985, (originally published by Thomas T. Cromwell, Co., New York, 1967.)

6. Lafflin, John, **Brassey's Battles**, Brassey's Defense Publishers, London, 1986.

7. Livermore, Thomas J., **Numbers and Losses in the Civil War in America 1861-1865**, Houghton, Mifflin and Company, Boston, 1900.

8. Jones, Archer, **The Art of War in the Western World**, University of Illinois Press, Urbana, 1987.

9. Dupuy, R. Ernest, and Trevor N. Dupuy, **The Encyclopedia of Military History**, Harper and Row, New York, 1977.

X. Arguments Against Lanchester Attrition Theory

X.A Introduction

In the preceding chapter, we reviewed the Lanchester validation efforts of Engel[1], Busse[2], Willard[3], and Fain[4]. The first two examined an individual battle/-campaign to test the merit of Lanchester Quadratic attrition while the latter two examined collections of battles to calculate attrition order. While the first efforts displayed some success, the latter resulted in an estimate of historical attrition order of ~ 2.5, substantially different from the value between 1 and 2 that we would have expected from Lanchester Theory.

Having introduced the serpent into Eden, it is appropriate now that we review the chief arguments against Lanchester Theory before we continue our examination of history.

One of the foremost critics of Lanchester Attrition Theory (LAT), as we have noted before, is Trevor Dupuy, whose Quantified Judgement Model (QJM) we will briefly review in the next chapter[5,6]. Dupuy's chief argument against Lanchester Theory is based on the Willard-Fain analysis.

Another critic is Joshua Epstein, an analyst who also has an alternative combat model, (which we shall also review in the next chapter,) presents three major problems with Lanchester Theory:[7,8]

(1) Why Withdraw?
(2) No Trading Space for Time, and
(3) No Diminishing Marginal Returns.

We shall address each in turn.

X.B. Why Withdraw?

Epstein states that LAT does not contain any feedback, that "not one of the equations can capture the effort of withdrawal - a response to attrition - on the rate of attrition itself." This is a simplification whose apparent truth masks the basic assumptions of Lanchester Theory.

Admittedly, Lanchester Theory does not contain any feedback mechanism to alter the attrition rates in a withdrawal. Lanchester Theory is not a theory of combat[a], but a theory of combat attrition; it should not be expected to automatically emulate these effects. This is not to say that the efforts of withdrawal on attrition cannot be

---

[a] For a somewhat consistent review, see Lepingwell, John W. R., "The Laws of Combat?", International Security, Summer 12. 89-134, 1987.

incorporated. These effects can be incorporated in the attrition rates via Bonder-Farrell Attrition Rate Theory (which we will discuss later in the book). Their effects do not occur automatically, however.

Technically, Epstein's first criticism is valid, but only if one misinterprets what Lanchester Attrition Theory is. If one correctly views it as only a theory of attrition, than this criticism is reduced to a statement of limitation.

X.C. No Trading Space for Time.

The second criticism is that the conclusion time predicted by Lanchester Theory does not reflect withdrawal or movement. Again, the criticism is true but is fallaciously based on the idea that Lanchester Theory is a general theory of combat. Thus, this criticism not only carries forward the same misunderstanding of previous criticism, but compounds it by misinterpreting the nature of the conclusion time.

The conclusion time is a mathematical convenience, a tool. It does not hold any relationship to actuality that his ever been demonstrated. If we accept the restriction that Lanchester Theory is a limited theory of attrition in combat only, then we must look elsewhere for models of the condition that initiate and terminate attrition! The conclusion time is most certainly <u>not</u> a model of that except under extraordinarily circumstances.

We may sketch a set of Lanchester differential equations that incorporate withdrawal and trading space for time. Assume that the force strengths of the two sides are functions of time and position, and that the attrition rates are (at least) functions of position. In this case, we may write a pair of attrition differential equations

$$\frac{d}{dt}A(\underline{r}_A,t) = -\alpha(\underline{r}_A,\underline{r}_B)B(\underline{r}_B,t), \tag{X.C-1}$$

and

$$\frac{d}{dt}B(\underline{r}_B,t) = -\beta(\underline{r}_B,\underline{r}_A)A(\underline{r}_A,t), \tag{X.C-2}$$

with the supplemental trajectory equations,

$$\underline{r}_A(t) = \underline{r}_A(0) + \int_0^t \underline{v}_A(t')\,dt', \tag{X.C-3}$$

and

$$\underline{r}_B(t) = \underline{r}_B(0) + \int_0^t \underline{v}_B(t')\, dt', \qquad \text{(X.C-4)}$$

where: A,B are the time and position dependent force strengths,
α, β are the position dependent attrition rates,
$\underline{r}_A$, $\underline{r}_B$ are the time dependent positions of the two forces, respectively,

and

$\underline{v}_A$, $\underline{v}_B$ are the time dependent velocities of the two forces, respectively.

We now divide the engagement into two parts in time. For $0 \le t \le t_1$, we take

$$\begin{aligned} \underline{v}_A(t) &= \underline{0}, \\ \underline{v}_B(t) &< \underline{0}, \end{aligned} \qquad \text{(X.C-5)}$$

in the sense that $|\underline{r}_A - \underline{r}_B|$ is decreasing, so that the A force is stationary (defending) and the B force is advancing (attacking). For $t_1 \le t \le t_2$, we change $\underline{v}_A$ so that

$$\underline{v}_A \le \underline{0}, \qquad \text{(X.C-6)}$$

so that the A (defending) force is now withdrawing, and probably

$$|\underline{v}_A| \ge |\underline{v}_B|,$$

so that the defending force is withdrawing faster than the B force is advancing. The effective zeros of the attrition rates with range separation would thus define t_2 at this separation as the close or end of the engagement. If we view the attrition rates as being given approximately (we will treat this in greater detail in the section of the book on attrition rate theory,) as

$$\alpha(\underline{r}_A, \underline{r}_B) \simeq \rho_B\, p_{kB}\, p_{LOS}(\underline{r}_A, \underline{r}_B), \qquad \text{(X.C-8)}$$

(and similarly for β,)

where: ρ_B is the Blue unit rate of fire,
p_{kB} is the Blue unit probability of kill per shot, and
p_{LOS} is the probability of Line Of Sight (LOS) between the two positions,

and

p_{LOS} is a symmetric function in its arguments, that is,

$$p_{LOS}(\underline{r}_A, \underline{r}_B) = p_{LOS}(\underline{r}_B, \underline{r}_A),$$

then the attrition rates will become zero when p_{LOS} becomes zero. (Distinctions between hull defilade and fully exposed target effects are, among other places,

contained in the p_k's.) This effectively closes the engagement at some time (which we are liberty to designate t_2,) given by the values of the positions.

This division of the movement into two regions has the effect of modeling withdrawal and trading space for time in the engagement using Lanchester Attrition Theory. We could, of course, embellish the model by using different attrition rates for the two divisions, and probably should, but this embellishment is unnecessary to demonstrate our point that these two criticisms of Epstein's can be treated with LAT. Admittedly, this model does not incorporate feedback, as Epstein's model of the next chapter does, but it does allow us to address the two criticisms.

X.D. No Diminishing Marginal Returns.

This criticism relates to the quadratic State Solution. The argument is that if one force is twice the other, then the second force must have an attrition rate four times the first's to force a stalemate. As Epstein points out, this is not born out by history, although history also displays that the assumptions of Lanchester Theory have been violated as well.

Simply put, it is possible for the state solution to apply, but only if the assumptions implicit in Lanchester Theory apply as well. As soon as one side becomes larger than the other, keeping all units in combat becomes problematical. As before, the criticism becomes limitation.

X.E. Back to History.

The criticism of Lanchester Attrition Theory thus come to be seen as turning on whether history will support the mathematics. As we have seen, Epstein's criticisms are fundamentally based in an expectation that Lanchester Theory is a general theory of combat, which it is not. Having eliminated the teeth of these criticisms of what is not, what remains is a question of whether the data of history will support Lanchester Theory?

Basically then, the question comes down to the analysis of Willard and Fain. As we have seen , these give rise to attrition orders of ~ 2.5 in their calculations, and substantially other values for our data bases. We must therefore examine these data basis in some more detail.

Before proceeding on this, we take time to examine the behavior of Willard's equation

$$\ln\left(\frac{b}{a}\right) = \ln\left(\frac{\beta}{\alpha}\right) + (n - 1) \ln\left(\frac{A_0}{B_0}\right), \tag{X.E-1}$$

under simulated conditions where we know n = 2. If we take a series of values for A_0, generate from distributions A, B_0., α and β, and calculate B from the state solutions, how likely are we to get n = 2? We conducted just such an experiment, and found, for this experiment, that we got an average attrition order of 1.86 with a standard deviation of 0.75. This is a very large standard deviation, but it indicates that an attrition order of 2.5 is not as strong a deviation from the desired value between 1.5 and 2 as we would expect. The noise in the method itself may be making the situation seem worse than it really is.

Next, we tried another attrition equation based on the differential equation itself.

$$\frac{dA}{dt} = -\alpha A^{2-n} B, \tag{X.E-2}$$

which we rewrite as

$$\frac{dA^n}{dt} = -\alpha n\, A\, B, \tag{X.E-3}$$

and integrate approximately as

$$A_0^n - A^n \simeq \frac{\alpha\ \tau\ n}{2} \left(A_0 B_0 + AB \right). \tag{X.E-4}$$

We expand the left hand side in the same manner as for Willard's equation and get

$$A_0^{n-1}\, a \simeq \frac{\alpha\ \tau}{2} \left(A_0 B_0 + AB \right), \tag{X.E-5}$$

where:

$$a \equiv A_0 - A = \Delta A. \tag{X.E-6}$$

We rewrite Equation (X.E-5) as

$$\ln\left(\frac{2\,a}{A_0B_0 + AB}\right) = \ln(\alpha\,\tau) + (1 - n)\,\ln(A_0), \qquad \text{(X.E-7)}$$

and curve fit it (and the similar equation for B.) This gives an average attrition order of 1.94 with a standard deviation of 0.08.

Next, we applied this fitting technique to the historical data introduced in the preceding chapter. The results are given in the Table.

Data Set	Attrition Order	Standard Error	Number of Battles
Nominal	1.883	0.065	108
Civil War	2.097	0.125	49
Osipov	2.124	0.139	38
Short	1.838	0.066	72
World War I	1.901	0.223	12

This gives an average attrition order of 1.942, which is in the region, between 1.5 and 2, sought by Willard and Fain. These results are much more consistent with what we would expect from Lanchester Theory, as demonstrated by the errors when compared to those calculated in the previous chapter.

Why should this fitting method be better than Willard's formula? If we make scatter plots of the independent versus dependent variables for Willard's equation, equation (X.E-1) (i.e. $\ln(a/b)$ versus $\ln(A_0/B_0)$) and for our approximately integrated differential equation, equation (X.E-7) (i.e. $\ln(2a/(A_0\,B_0 + A\,B))$ versus $\ln(A_0)$ and $\ln(2b/(A_0\,B_0 + A\,B))$ versus $\ln(B_0)$) we see an apparent reason. We present these for the Nominal Length Battles data set in Figures (IX.D.1) and (IX.D.2) respectively. Imagine how you woulr draw a straight line through these data (which is what the linear regression will do.) My intuitive guess for doing this in both figures would draw the line running from upper left to lower right. This intuition is completely wrong for the result we desire from Willard's formula since for $n = 2$ we would want a slope of one! The intuition is right for the appriximately integrated differential equation since for $n = 2$ we would want a slope of minus one. Further, notice the high degree of scatter in figure (IX.D.1) as compared to figure (IX.D.2). This tighter pattern in the second figure contributes to a better fit in the linear regression. Thus, while we would expect a better fit from Willard's formula, since it is based on exact integration and

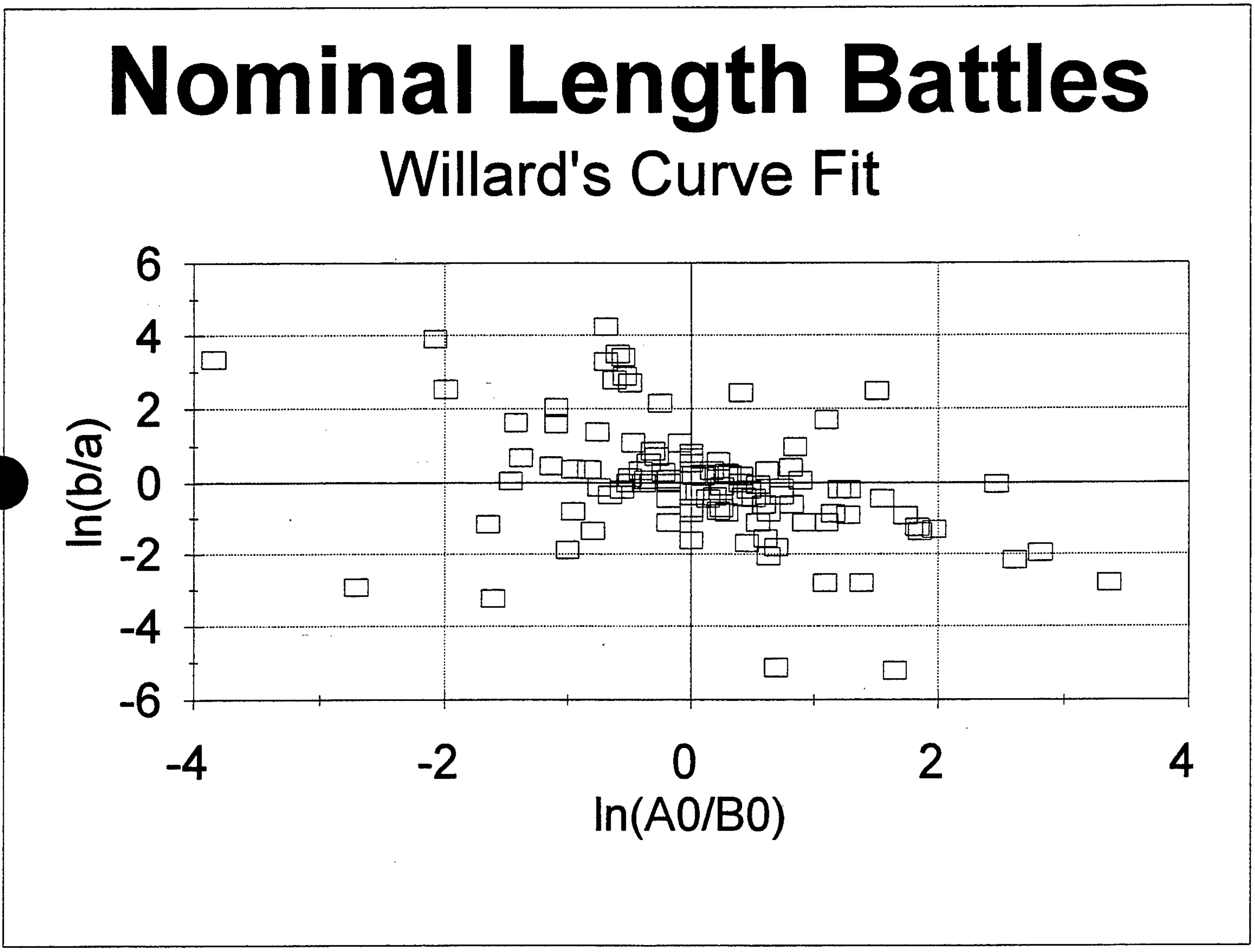
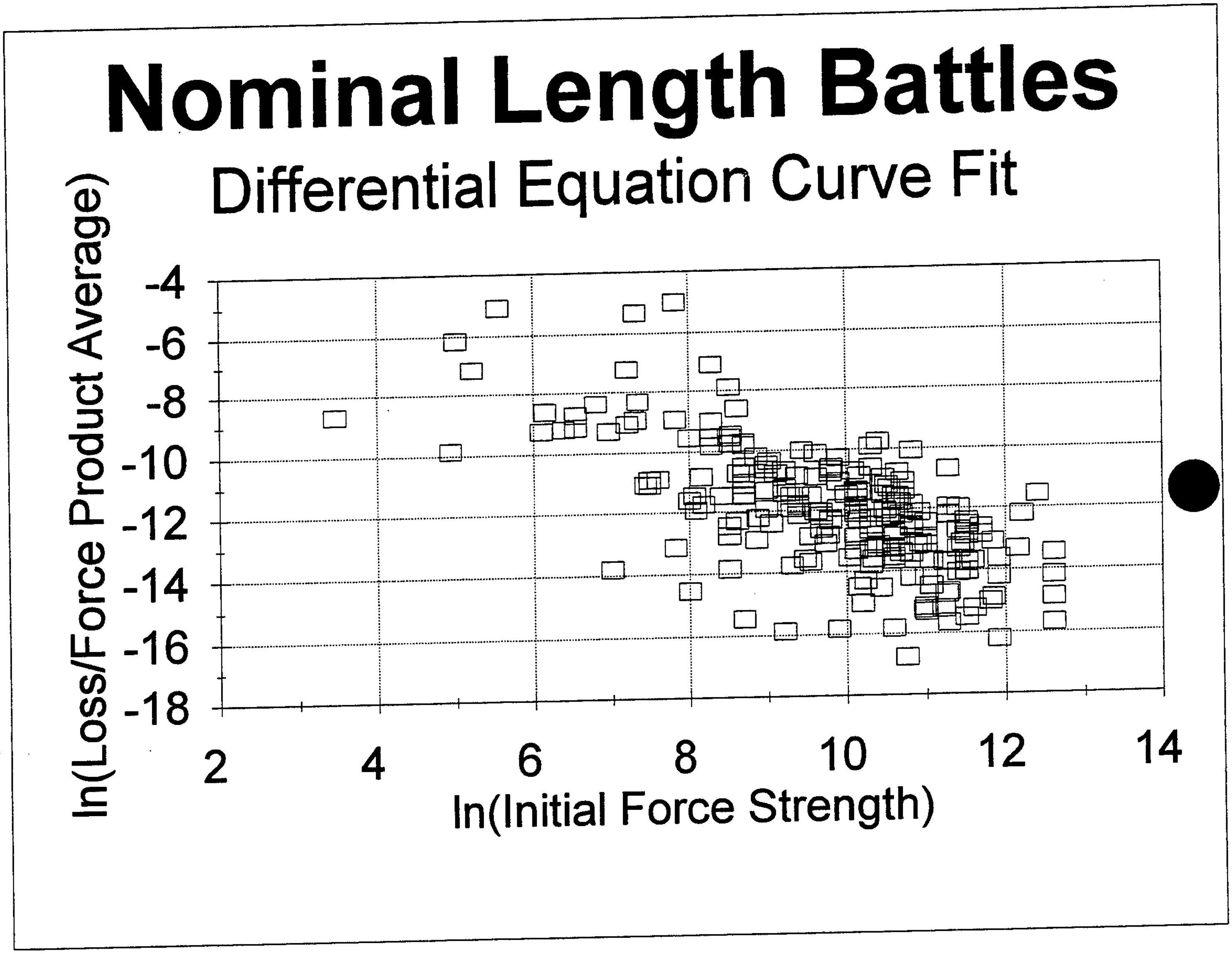

Nominal Length Battles
Willard's Curve Fit
6
4
2
0
-2
-4
-6
ln(b/a)
-4
-2
0
2
4
ln(A0/B0)

Figure IX.D.1



Nominal Length Battles
Differential Equation Curve Fit
ln(Loss/Force Product Average)
-4
-6
-8
-10
-12
-14
-16
-18
2
4
6
8
10
12
14
ln(Initial Force Strength)

Figure IX.D.2

incorporates the same approximation of loss representation as we used in the approximately integrated differential equation, it does not provide us with as compact a set of data as the latter equation. Despite the greater degree of approximation introduced in the approximately integrated differential equation approach, it therefore appears reasonable to accept it as a better estimator of attrition order than Willard's equation.

There is another potential source of error that we should also recognize. In Willard's equation, the intercept of the fitted line is the logartithm of the ratio of the attrition rates, while in the approximately integrated differential equation, the intercept is the logarithm of the attrition rate. If we view the attrition rates during as random variables then the approximately integrated differential equation will fit the intercept (approximately) to the mean of the distribution. This is not the case with Willard's equation. An initial view would lead us to believe that the intercept for Willard's equation should be approximately zero if the two attrition rates are drawn from the same distribution and we are calculating the ratio of the mean of that distribution to the mean of the same distibution. This is not the case, however. What we are calculating is the mean of the logarithms of a set of sample draws from the distribution. The distribution of these ratios is considerably less well behaved than the distribution of the attrition rates. Any skewness in the distribution about the mean is magnified and this can lead to greater distortion in the data and the resulting fit.

REFERENCES

1. Engel, J. H., "A Verification of Lanchester's Law", Journal of Operations Research, 2 163-171, 1954.

2. Busse, James J., "An Attempt to Verify Lanchester's Equations", Paper 9-1, pp 587-597, in Benjamin Avi-Itzhak, ed., **Developments in Operations Research**, Vol. 2, Gordon and Breach, New York, 1969.

3. Willard, Daniel A., "Lanchester as a Force in History: An Analysis of Land Battles of the Years 1618-1905", McLean, VA, 1962, as cited in Dupuy, **Numbers, Predictions and War.**

4. Fain, Janice B., "The Lanchester Equations and Historical Warfare: An Analysis of Sixty World War II Land Engagements", Arlington, VA 1975, as cited in Dupuy, **Numbers, Predictions and War.**

5. Dupuy, T. N., COL U.S. Army, Ret., **Numbers, Predictions and War,** Bobbs-Merrill Co., Inc., New York, 1979.

6. Dupuy, Trevor N., COL, U.S. Army, Ret., **Attrition: Forecasting Battle Casualties and Equipment Losses in Modern War,** Hero Books, Fairfax, VA, 1990.

7. Epstein, Joshua M., **The Calculus of Conventional War: Dynamic Analysis without Lanchester Theory,** Studies in Defense Policy Series, The Brookings Institution, Washington, 1985.

8. Epstein, Joshua M., **Strategy & Force Planning: The Case of the Persian Gulf**, The Brookings Institution, Washington, 1987.

XI. Two Alternatives to Lanchester

XI.A. Introduction.

In the previous chapter, we reviewed some of the primary criticisms of Lanchester Attrition Theory, noting therein that two of the critics: Trevor Dupuy and Joshua Epstein had advanced combat models of their own as alternatives to (or improvements over) the Lanchester model. The purpose of this chapter is to briefly review each of these models to provide a basis of comparison between the bulk of the book and some of the alternatives. This review will be sketchy and cannot begin to do justice to the efforts of these two workers in developing their models. I apologize here and now for errors and omissions of explanation caused by my lack of understanding.

XI.B. The Quantified Judgement Model

COL Trevor Dupuy is a well-know military historian; his Quantified Judgement Model (QJM) reflects that, being based on Clausewitz's "Law of Numbers"

> *"If we ... strip the engagement of all the variables arising from its purpose and circumstances, and disregard (or strip out) the fighting value of the troops involved (which is a given quantity), we are left with the base concept of the engagement ... in which the only distinguishing factor is the number of troops on either side.*
> *These numbers, therefore, will determine victory (and are) the most important factor in the outcome of an engagement...*
> *This ... would hold true for Greeks and Persians, for Englishmen and Maharattas, for Frenchmen and Germans"*[1]

The basic description of the QJM is **Numbers, Predictions, and War**[2] , updated by **Attrition: Forecasting Battle Casualties and Equipment Losses in Modern War**[3]. The student must be prepared for diligent inspection when studying these texts. COL Dupuy is not a mathematician, his formulas are plagued with apparent inconsistencies and errors. Parameters that change with tactical era are not always vigorously identified, and some alternative definitions may border on being contradictory. Many of the formula are derived from historical data so they may not appear obviously logical, or agree with similar formula obtained by other means. We must keep in mind that there are two sets of messages here: the historical and the mathematical; and not allow the complexities and mistakes in the latter to destroy our learning from the former.

The fundamental relationship of the QJM, taken from Clausewitz's Law of

Numbers is the combat power of a side, defined by

$$P = S\ V\ C_{ev}, \tag{XI.B-1}$$

where: S = Force Strength (which is different from the Lanchestrian),
V = Operational Variables, and
C_{ev} = Combat Effectiveness Value.

The Combat Effectiveness Value is judgmental factor, related by Dupuy to nationality and generalship, that modifies the combat power. Dupuy admits that "there is as yet no scientific way to forecast C_{ev}'s."

The force strength is the summation over all the elements of the force of

$$S = \sum_{allelements} W_i\, r_i\, h_i\, z_i\, w_i\ ,$$

where: W_i = Operational Lethality Index,
r_i = terrain factor,
h_i = weather factor,
z_i = sensor factor, and
w_i = air superiority factor.

The Operational Lethality Index is calculated as the quotient of the Theoretical Lethality Index and the Dispersion Index (the area in km^2 occupied by a force of 100,000) which is a function of tactical era and situation. It represents a density of forces on the battlefield. The Theoretical Lethality Index formula depends on the type of weapon, and includes the technical characteristics of the weapon including lethality, accuracy, and vulnerability. Thus the Operational Lethality Index can be thought of as being proportional to the (Lanchestrian) Force Strength times the attrition rate.

The Operational Variables V is a product of several factors including mobility, leadership, training, morale, logistics, military posture, terrain, weather, season, and vulnerability. This variable incorporates the type of engagement, and distinguishes attacker from defender.

Dupuy predicts battle outcome based on two quantities: the Combat Power - the larger should be victorious, and the result formula

$$R = M_f + E_{sp} + E_{cas}, \tag{XI.B-3}$$

where: M_f = mission accomplishment,
E_{sp} = spatial effectiveness, and
E_{cas} = casualty effectiveness,

which should also be larger for the victor. The mission accomplishment is judgmental. Spatial effectiveness is calculated from both sides' force strength and advance rates. Casualty effectiveness is calculated from casualty predication formula.

In summary, the QJM is a general model of combat based on historical development. It is considerably more complicated than the basic Lanchester model. Its chief drawbacks are its complexity, often contradictory formalism, its inherently judgmental components, and its lack of detailed consideration in a scientific sense. It is also a relatively fragile model. It must be used in an "all or nothing " form, it has no scientific basis for introducing new technology for consideration, and it may be too uniquely embedded in the warfare database of its origin.

XI.C. The Epstein Model

This model, considerably similar in form than the QJM, is documented in **The Calculus of Conventional War**[4] and **Strategy and Force Planning**[5]. The model is structurally somewhat similar to Lanchester attrition. It has a finite difference form with time increments of days.

The attacker ground lethality evolution equation is

$$A_g(t) = A_g(t-1) - \alpha(t-1)\,A_g(t-1) - C_{asD}(t-1), \qquad \text{(XI.C-1)}$$

where: α(t-1) = attacker's total ground lethality attrition rate per day (on (0,1))
C_{asD}(t) = attacker's ground lethality killed on day t by defender's close air support.

Note that α is the defender's attrition rate in Lanchester terms, expressed as a fraction of the attacker's strength. In this model, instead of force strengths in numbers, the operant quantities are ground (and air) lethality. For homogeneous aggregation, it would seem that total force strength and force ground lethality are approximately linearly related by a constant?

The defender ground lethality evolution equation is:

$$D_g(t) = D_g(t-1) - \frac{\alpha(t-1)}{\rho}\,A_g(t-1) - C_{asA}(t-1), \qquad \text{(XI.C-2)}$$

where: C_{asA}(t) = defender's ground lethality killed on day t by attacker's close air support, and
ρ = attacker's ground lethality killed per defender's ground lethality killed (average ground to ground exchange ratio.)

The ground lethality rate is given by
where: α_g = attacher's ground-prosecution rate,

$$\alpha(t) = \alpha_g(t)\left(1 - \frac{w(t)}{w_{MAX}(t)}\right), \qquad \text{(XI.C-3)}$$

$w(t)$ = defender's rate of withdrawal, and
$w_{MAX}(t)$ = defender's maximum rate of withdrawal.

The withdrawal rate has an evolution equation,

$$\begin{aligned} w(t) &= 0, \alpha_d(t-1) \le \alpha_{dT}, \\ &= w(t-1) + \frac{w_{MAX} - w(t-1)}{1 - \alpha_{dT}}(\alpha_d(t) - \alpha_{dT}), \alpha_d(t-1) > \alpha_{dT}, \end{aligned} \qquad \text{(XI.C-4)}$$

where:

$$\alpha_d(t) = \frac{D_g(t) - D_g(t-1)}{D_g(t)}, \qquad \text{(XI.C-5)}$$

and α_{dT} = defender's threshold attrition rate.

The close air support aircraft surviving on with day t have the forms,

$$D_a(t) = D_a(1)\,(1 - \alpha_{da})^{S_d(t-1)}, \qquad \text{(XI.C-6)}$$

for the defender,and

$$A_a(t) = A_a(1)\,(1 - \alpha_{aa})^{S_a(t-1)}, \qquad \text{(XI.C-7)}$$

where: α_{da}, α_{aa} = defender, attacker aircraft attrition rate per sorte, on (0,1),
and S_d, S_a = defender, attacker sorte rate.

The ground lethality killed by close air support aircraft are:

$$C_{asD}(t) = \frac{L}{V} D_a(t)\, K_d \left[\frac{1 - (1 - \alpha_{da})^{S_d+1}}{\alpha_{da}} - 1 \right], \qquad \text{(XI.C-8)}$$

and

$$C_{asA}(t) = \frac{L}{V} A_a(t)\, K_a \left[\frac{1 - (1 - \alpha_{aa})^{S_a+1}}{\alpha_{aa}} - 1 \right], \qquad \text{(XI.C-9)}$$

where: L is the number of "lethality points" per division equivalent,
V is the number of armored fighting vehicles per division equivalent, and

K_d, K_a are the defender, attacker Close Air Support kills of enemy AFVs.

Epstein claims that this model incorporates feedback on both attacker and defender sides that moderates the battle. It clearly has feedback and while we have not exercised the model sufficiently to completely establish the extent of the feedback, it is already there.

This model is more general than what we have seen so far in Lanchester theory. We do not believe it is as general as Lanchester theory in terms of admitting different aggregations' of forces. It has not been subjected to the validation efforts that the Lanchester Theory has.

Perhaps the most difficult thing about this model however is an apparent tautology. The quantity ρ is defined as the ratio of attacker ground lethality killed to defender ground lethality killed. Thus

$$\rho = \frac{\Delta A_g(t)}{\Delta D_g(t)}, \tag{XI.C-10}$$

where we have reintroduced the Δ notation.

If we substitute equation (XI.C-10) into equation (XI.C-2) , we get, (changing to the Δ notation we are used to,)

$$\begin{aligned} \Delta D_g(t) &= -\frac{\alpha(t-1)}{\rho} A_g(t-1) - C_{asA}(t-1) , \\ &= -\alpha(t-1) \frac{\Delta D_g(t)}{\Delta A_g(t)} A_g(t-1) - C_{asA}(t-1) , \end{aligned} \tag{XI.C-11}$$

which we may simplify as

$$\Delta A_g(t) = -\alpha(t-1) A_g(t-1) - C_{asA}(t-1) \frac{\Delta A_g(T)}{\Delta D_g(t)} . \tag{XI.C-12}$$

This is not the same as equation (XI.C-1) unless the last term reduces to $C_{asD}(t-1)$.

This is the fundamental problem with the Epstein model. While it does have the feedback that allows trading space for time, it has no accommodation for calculating the attrition rates on the firm scientific basis that Lanchester theory does. It is useful for studying that feedback correlation, but not for evaluating the interplay of different force components and new weapons technology on the battlefield.

XI.D. Conclusion.

We have examined both the QJM of Dupuy and the Epstein model. Both are more general models of combat than Lanchester Attrition Theory, and both are somewhat more specialized models which do not have the generality to allow for the evaluation of force components and technology. They provide useful insights into combat processes, but their greater "correctness" reduces their generality compared to the mechanics of Lanchester Attrition Theory.

REFERENCES

1. Clausewitz, Carl von, **On War**, Book 3, Chapter 8.

2. Dupuy, T. N., COL, U.S. Army, Ret., **Numbers, Predictions and War**, Bobbs-Merrill Co., Inc., New York, 1979.

3. Dupuy, Trevor N., COL, U.S. Army, Ret., **Attrition: Forecasting Battle Casualties and Equipment Losses in Modern War**, Hero Books, Fairfax, VA, 1990.

4. Epstein, Joshua M., **The Calculus of Conventional War: Dynamical Analysis without Lanchester Theory**, Studies in Defense Policy, The Brookings Institution, Washington, 1985.

5. Epstein, Joshua M., **Strategy and Force Planning: The Case of the Persian Gulf**, The Brookings Institution, Washington, 1987.

XII. The Recent Unpleasantness

XII.A. Introduction

As a child growing up in the South, I frequently heard little old ladies, at least my grandmother's age, using this term to refer to what we now seem to have settled into calling the Civil War.[a] These ladies had learned the term from their grandmothers or even great-grandmothers, who may have experienced the shortages and emotions of the war itself, but assuredly had experienced the frustration and agony of Reconstruction. To these ladies, <u>the</u> War was not a matter of military effort, but the impact of the war on their daily lives.

The Civil War was much closer to the people in those days. The rite of passage of becoming a teenager, and the Centennial of the war are irrevocably linked periods for me. The social acceptability of beards turned about within a matter of months and fashion retreated to hoop skirts and frock coats on such a multitude of occasions as to become the norm. No cemetery, less than a half century old, did not contain markers commemorating that the penultimate activity of the man interred thereunder was to serve in the war; his life and accomplishments afterwards being too mundane and colorless for memory

While no Southron community has forgotten the war, most Northern communities have. Except for those whose economy is dominated by a battlefield or a war hero's home, the only Northern community with memory, that I have found, is Carlisle, PA, and their memory is largely limited to Stuart's raid of '63. Still, there are pockets of interest, evidenced by civilian commemorative units who reenact battles even in places where they never occurred, such as California and New Jersey.

Modern liberals may decry this seeming worship of a war whose *raison d'etre*, in their minds, was largely racial. Still, the Civil War is also a matter of serious professional study, as evidenced by the fact that one in five of history books written in this country deal with the war.

Well may we study the Civil War. Its influence on our society has been more profound than any war except the American Revolution, and its evolvement of the general populace was much more widespread. The span of its battles and engagements is enormously greater, with some authorities claiming the number exceeds five thousand.

---

[a] Also known as the War Between The States, the War of Southron Independence, the War of Northern Aggression, and the Second American Revolution, to cite only a few, clearly with a Southern proclivity.

In our current age, there are still many lessons to be learned; many insights to be drawn from the war. In our current environment of high and rising technology, and the reduction of the sizes of the armed services in the demise of the Soviet Union, there is considerable information. By studying the war, we may hope to learn how armies adopt technology and develop tactics and doctrine. The advent of railroads, breech loading guns, rifling, and connodial bullets were no less military innovations in their day than electronics and guided weapons are in ours.

In some cases, the technology was successfully digested into doctrine, as with railroads. Other technologies were not so successfully adopted. A case in point were longer range, more accurate, and higher rate of fire infantry weapons. This failure has been ascribed as the cause of high attrition rates, especially for the Confederacy.[1,2] We shall examine this in subsequent sections.

Another area of interest is the rapid expansion of the armed forces from small professional cadres to large volunteer(?) establishments, and the development of organizational methods for intermixing and managing state (i.e. National Guard or Reserve) units with national (i.e. Active Component) units. While this was not the only time this has occurred in American history, it was the first time on the scale of a major national conflict rivaling the first and second world wars. Notable in this process (among other examples,) was how the armies were able to foster the growth and advancement of civilians who proved exceptional soldiers, as evidenced by men such as John Singleton Mosby and Joshua Lawrence Chamberlain, while surviving and culling inept, often political, appointed to high rank without benefit of training or extensive experience.

Our interest here is to examine the Civil War as it provides evidence for and against Lanchester Attrition Theory. We are not primarily concerned with the political and many of the military aspects of the war. The battles and engagements are a source of data and because of the extensive study and documentation of the war, they are an almost unique source of such data.

In examining the Civil War, we shall approximately follow the outline and much of the content of Weiss' seminal article on the war.[3] We shall not attempt to completely reproduce that article here as it is readily available through professional and collegiate libraries. The student should study the article to fully appreciate nuances and differences of interpretation between myself and Weiss. Our treatment here will be somewhat different, as dictated by the needs of a textbook rather than a research paper, and reflecting some of the greater advantages now available in digital computers for data analysis. For the peace of mind of the student, especially the one not particularly mathematically or computer oriented, I will emphasize that most of the analysis presented here has been accomplished with a standard personal

computer spreadsheet program,[b] with only occasional use of a statistics program.[c] I wish to emphasize that I resort to the latter out of a desire for convenience and simplicity; all of the calculations could have been done with the spreadsheet program but with greater effort. (I would have had to put in the formulas explicitly.)

XII.B. Data and Statistics

In a footnote to the introduction of his paper, Weiss states that the work presented is "the result of the author's hobby." Be this as it may, it is clear from the article that hobbies do not have to be amateurish. This article amply demonstrates that the Victorian custom of scientific research as hobby did not die with that monarch. This paper is an excellent example of small science (i.e. funded not by some government or philanthropic organization, or conducted in some enormous university research laboratory,) at its best.

Weiss briefly describes the previous efforts to apply Lanchester Theory to historical data: Engel and Willard; described in earlier chapters. Weiss notes that "Willard's conclusion that ""there is little value in a simple version of Lanchester's equations as a predictive tool, where the only known quantities are initial strengths"" is both a discouragement and a challenge." We have examined, in the preceding chapter, an alternative approach than Willard's to estimating attrition order, and we shall continue that analysis in this chapter.

Weiss goes on to note, as we have, the relative wealth of data on the Civil War, its importance as a precursor of modern mechanized warfare, and the evidence arguing against "computerized war". Perhaps equally important, beyond the wealth of data, is the temporal and geographic compactness of these data. While this detracts from drawing general conclusions about WAR from the data, it also makes any trends and correlations easier to accept. Further, it provides a compact set of data that may provide insights that can be tested against the more general data sets, but which may have been hidden in their generality. (We do concede considerable tactical evolution during the progress of the war, as noted by military historians.[4])

---

[b] I started out using Quattro Pro (r) from Borland, and upon shifting to a Windows(r) environment, changed first to Excel(r), and then to Quattro Pro for Windows(r). This is not an endorsement for any spreadsheet program or any coding house. I merely want to emphasize that any personal computer, Macintosh, or workstation spreadsheet program will support most of the analysis and graphical presentation needed.

[c] While most spreadsheet programs will perform linear regression, they will not perform more elaborate statistical tests and calculations, such as correlations. There are a variety of programs that may be used to do these calculations, but I have used STASTIX(r) from Analytical Software for this effort because I was able to buy a copy at an Operations Research Society of America Convention. Thus, my selection was based on its ease.

Weiss draws his data from three sources: Phister[5], Livermore,[6] and Bodart.[7] The latter source is common in use with Willard. Of the three, only Livermore has been available to us. For convenience, we reproduce these data in Table (XII.B.1). As before, we include only those battles and engagement for which both initial and final force strengths of both sides are recorded; durations in days are also available. Weiss states that Livermore is the most meticulous of the three, that he designates winner and loser, and he distinguishes between assaults on fortified lines and other battles. While Livermore lists 64 battles, only 49 have complete data.

He concludes after initial analysis that the cumulative losses on both sides are approximately equal when summed over the whole war. Further, casualties occur at approximately a constant rate for the Confederacy, with 1863 and 1865 being low rate years, and 1864 being a high rate year for the Union, obviously reflecting Grant's strategy.[8] He also gives the distribution of number of Union battles by loss and shows that the number of battles is piecewise distributed by loss to the -2/3, -1, and -3/2 power. (Weiss presents figures showing the cumulative loss and the loss distribution that we do not reproduce here.)

Weiss also examines other time behaviors and distributions. We present equivalent figures to his for our data set. In Figures (XII.B.1) - (XII.B.4), we show scatter plots of Union initial and final strengths, killed, and wounded plotted versus starting date of each battle or engagement. The equivalent data for the Confederate side is given in Figures (XII.B.5) - (XII.B.8).

In Figure (XII.B.9), we present a scatter plot of battle duration versus date. Figure (XII.B.10) shows Confederate:Union initial force strength ratios (i.e. C_0/U_0) versus date.[d] Figure (XII.B.11) shows Confederate:Union loss ratios (i.e. $\Delta C/\Delta U$), versus date.[e] The student is free to examine these figures in search of pattern or trend; I am unable to find one. This is consistent with our earlier investigations. We note that while Confederate losses become selectively more extreme in the latter part of the war, as noted by historians, attributing this to the greater tactical and operational sophistication of Union leaders, and the greater urgency of staving off thrusts toward Richmond. Regardless, Confederate losses (absolutely,) are less than Union losses in more than 57% of the battles in our data set.

---

[d] Weiss uses several force ratios in his paper. We present a table of some of the more widely accepted and used force ratios in Appendix E, and discuss them in greater detail in a later chapter.

[e] The initial force strength ratio is oftem called the initial force ratio, I_0; the ratio C/U is often called the remaining force ratio (often R, but not to be confused with Weiss' R in a later section of this chapter;) and the ratio of losses $\Delta C/\Delta U$, is the Loss Exchange Ratio, L_{ER}.

Battle	Date	Union				Confederate				Duration
		Start	Finish	Killed	Wounded	Start	Finish	Killed	Wounded	
Bull Run	21-Jul-61	28452	26960	481	1011	32232	30263	387	1582	1
Wilson's Creek	12-Aug-61	5400	4456	223	721	11600	10443	257	900	1
Fort Donelson	12-Feb-62	27000	24392	500	2108	21000	19000	2000		4
Pea Ridge Arkansas	07-Mar-62	11250	10067	203	980	14000	13400	600		1
Shiloh	06-Apr-62	62682	52520	1754	8408	40335	30600	1723	8012	1
Williamsburg	04-May-62	40768	38902	456	1410	31823	30253	1570		1
Fair Oaks	31-May-62	41797	37413	790	3594	41816	36087	980	4749	1
Mechanicsville	26-Jun-62	15631	15375	49	207	16356	14872	1484		1
Gaines's Mill	27-Jun-62	34214	30213	894	3107	57018	48267	8751		1
Peach Orchard,	29-Jun-62	83345	78376	724	4245	86748	78146	8602		2
Seven Day's Battles	25-Jun-62	91169	81373	1734	8062	95481	75742	3478	16261	6
Cedar Mountain	09-Aug-62	8030	6271	314	1445	16868	15530	231	1107	1
Manassas & Chantilly	27-Aug-62	75696	65600	1724	8372	48527	39419	1481	7627	6
Richmond KY	29-Aug-62	6500	5450	206	844	6850	6400	78	372	1
South Mountain	14-Sep-62	28480	26752	325	1403	17852	15967	325	1560	1
Antietam	16-Sep-62	75316	63659	2108	9549	51844	40120	2700	9024	1
Corinth	03-Oct-62	21147	18951	355	1841	22000	19530	473	1997	1
Perryville	08-Oct-62	36940	33244	845	2851	16000	12855	510	2635	1
Prairie Grove Ark	07-Dec-62	10000	9012	175	813	10000	9019	164	817	1
Fredericksburg	13-Dec-62	106007	95123	1284	9600	72497	67841	595	4061	1
Chickasaw Bayou	27-Dec-62	30720	29507	208	1005	13792	13595	63	134	2
Stone's River	31-Dec-62	41400	32180	1677	7543	34732	25493	1294	7945	1
Arkansas Post	11-Jan-63	28944	27912	134	898	4564	4455	28	81	1
Chancellorsville	01-May-63	97382	86213	1575	9594	57352	46606	1665	9081	3
Champion Hill	16-May-62	29373	27119	410	1844	20000	17819	381	1800	1
Port Hudson Assault 1	27-May-63	13000	11162	293	1545	4192	3957	235		1
Port Hudson Assault 2	14-Jun-63	6000	4396	203	1401	3487	3440	22	25	1
Gettysburg	01-Jul-63	83289	65605	3155	14529	75054	52416	3903	18735	3
Fort Wagner Assault	18-Jul-63	5264	4138	246	880	1785	1616	36	133	1
Chickamauga	19-Sep-62	58222	46809	1657	9756	66326	49340	2312	14674	1
Chattanooga	23-Nov-63	56359	50884	753	4722	46165	43644	361	2160	2
Mine Run	27-Nov-63	69643	68371	173	1099	44426	43746	110	570	4
Olustee FLA	20-Feb-64	5115	3760	203	1152	5200	4266	93	841	1
Pleasant Hill	09-Apr-64	12647	11653	150	844	14300	13300	1000		1
Wilderness	05-May-64	101895	87612	2246	12037	61025	53275	7750		2
Drewry's Bluff	12-May-64	15800	13030	390	2380	18025	15729	355	1941	4
The Mine	30-Jul-64	20708	17844	2864		11466	10847	619		1
Weldon Railroad	18-Aug-64	20289	18986	198	1105	14787	13587	1200		3
Atlanta	03-May-64	110123	99595	10528		66089	56902	9187		28

Table B.1. Civil War Battles

Battle	Date	Union				Confederate				Duration
		Start	Finish	Killed	Wounded	Start	Finish	Killed	Wounded	
Kenesaw Mountain	27-Jun-64	16225	14226	1999		17733	17463	270		1
Tupelo	13-Jul-64	14000	13364	77	559	6600	5274	210	1116	2
Peach-Tree Creek	20-Jul-64	21655	20055	1600		18832	16332	2500		1
Atlanta	22-Aug-64	30477	28488	430	1559	36934	29934	7000		1
Atlanta	28-Jul-64	13226	12667	559		18450	14350	4100		1
Jonesborough GA	31-Aug-64	14170	13991	179	0	23811	22086	1725		1
Winchester	19-Sep-64	37771	33091	697	3983	17103	15000	276	1827	1
Cedar Creek	19-Oct-64	30829	26755	644	3430	18410	16550	320	1540	1
Franklin	30-Nov-64	27939	26717	189	1033	26897	21347	1750	3800	1
Bentonville	19-Mar-65	16127	15194	139	794	16895	15387	195	1313	1

Table [illegible]1. Civil War Battles

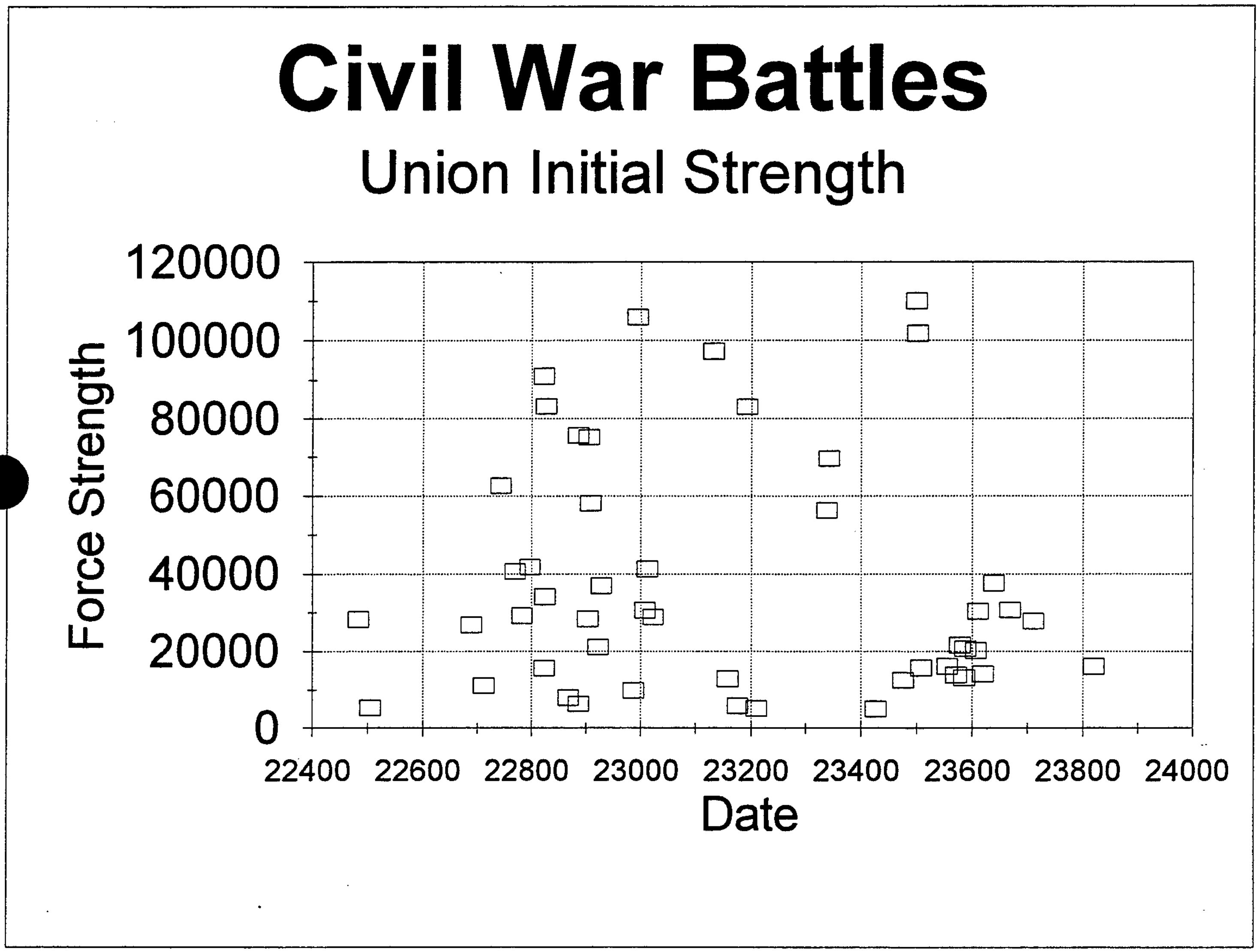


Figure XII.B.1

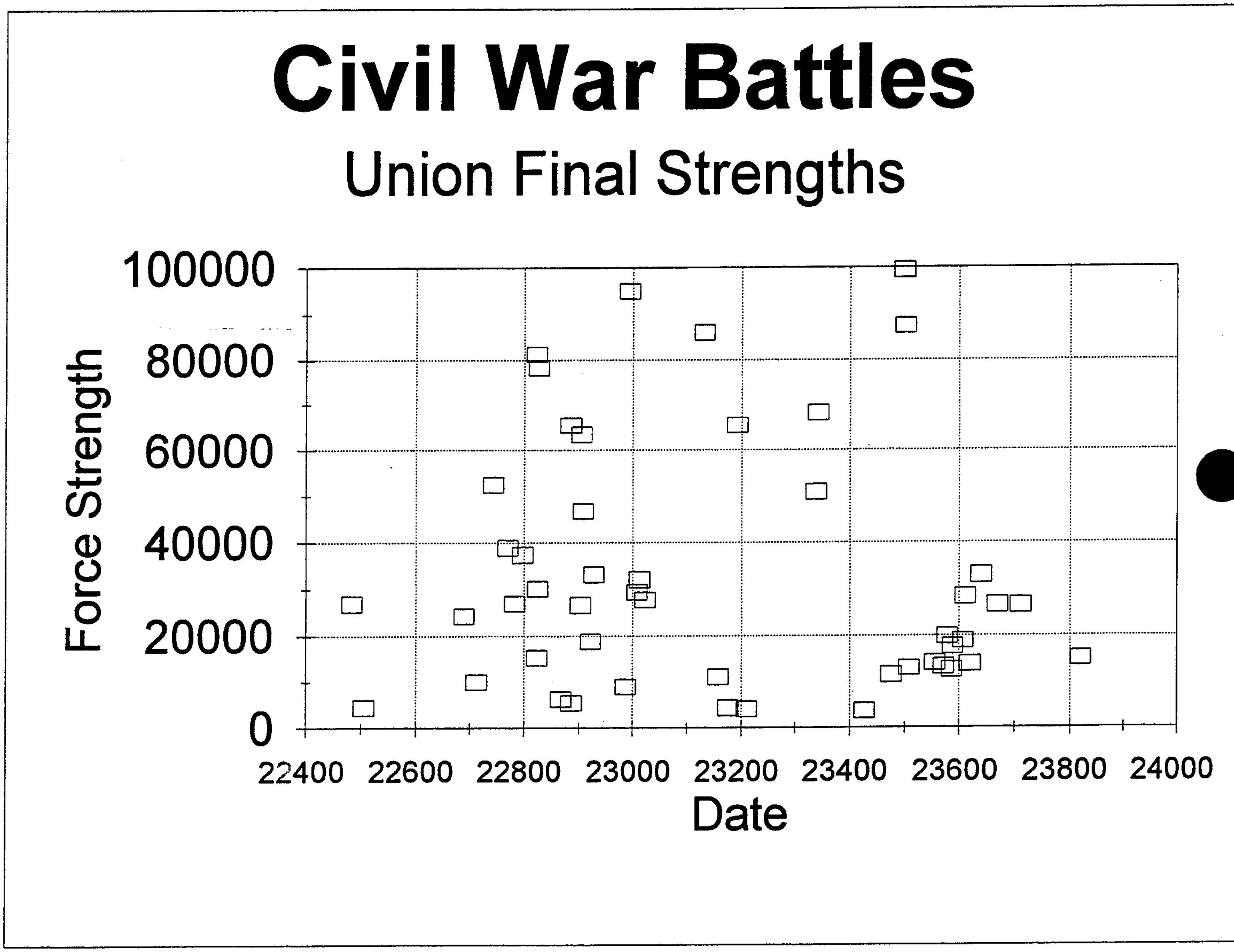


Figure XII.B.2

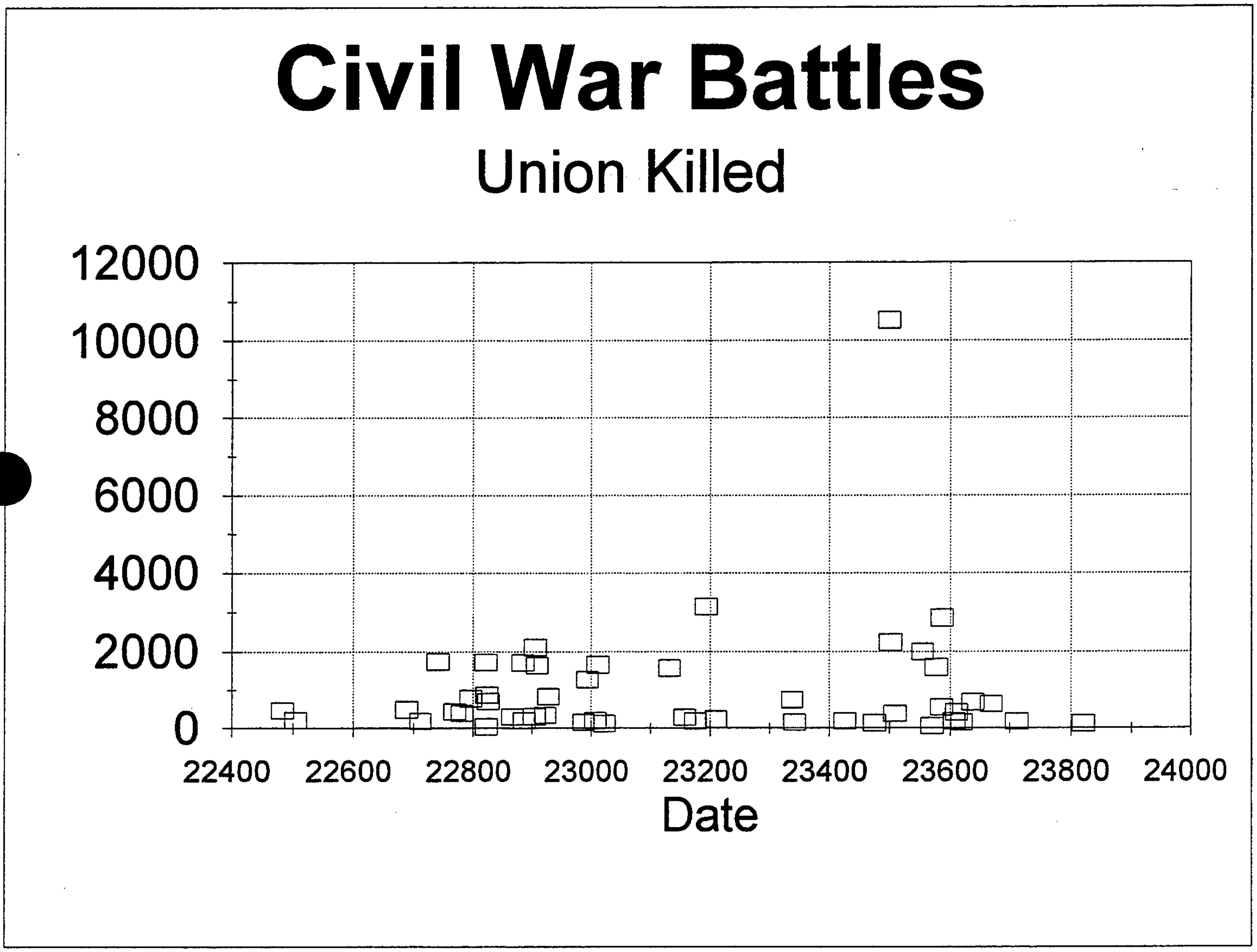


Figure XII.B.3

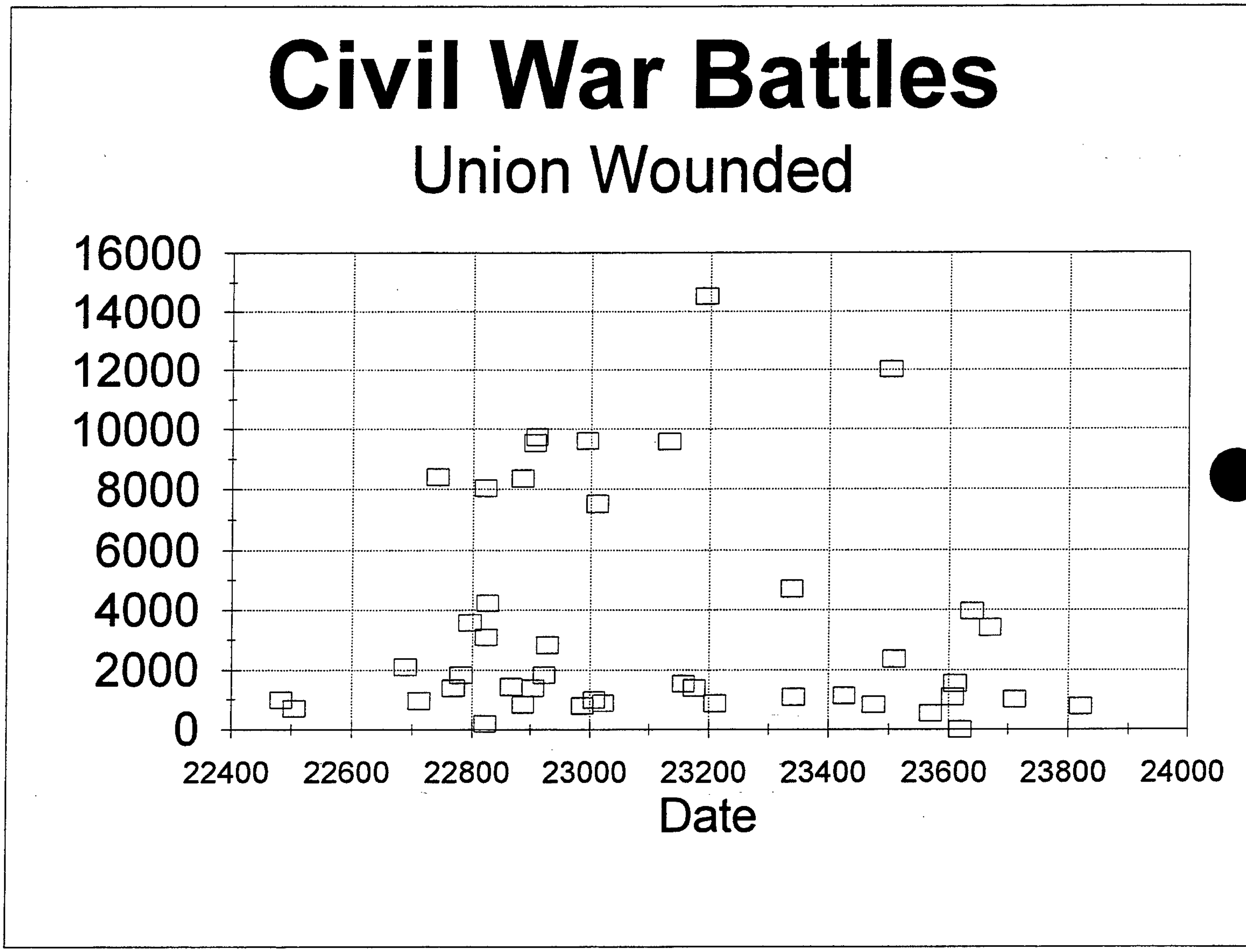


Figure XII.B.4

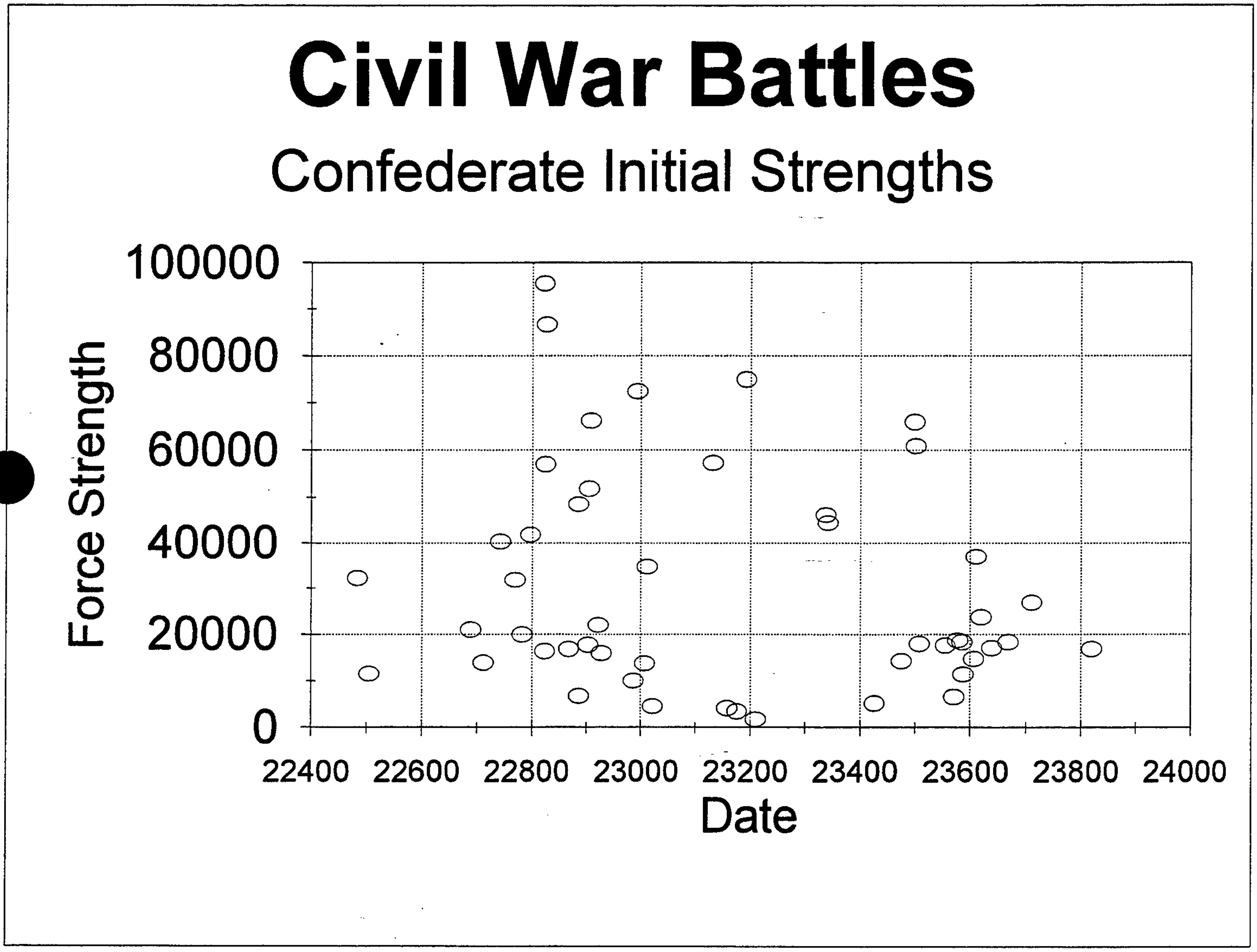


Figure XII.B.5

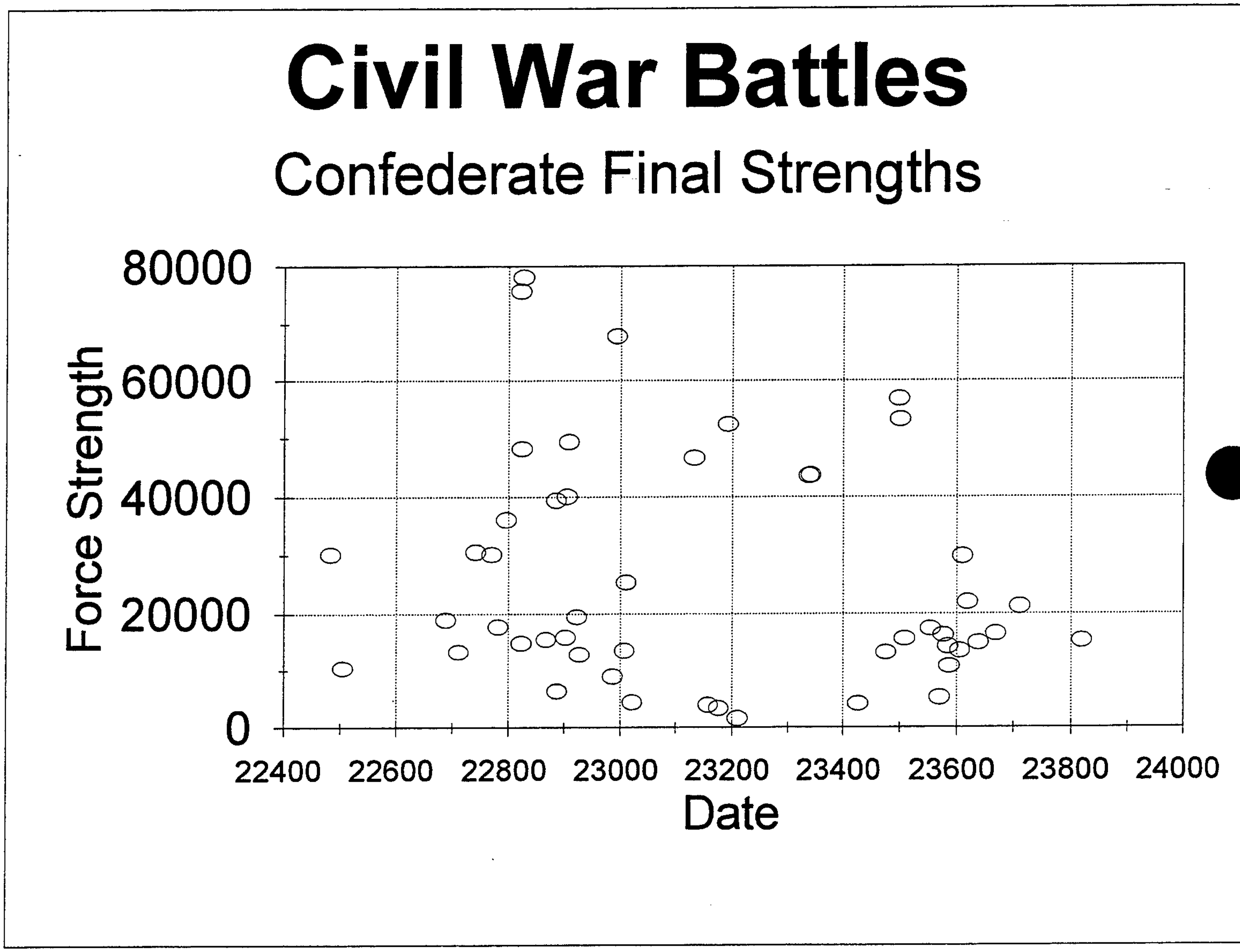


Figure XII.B.6

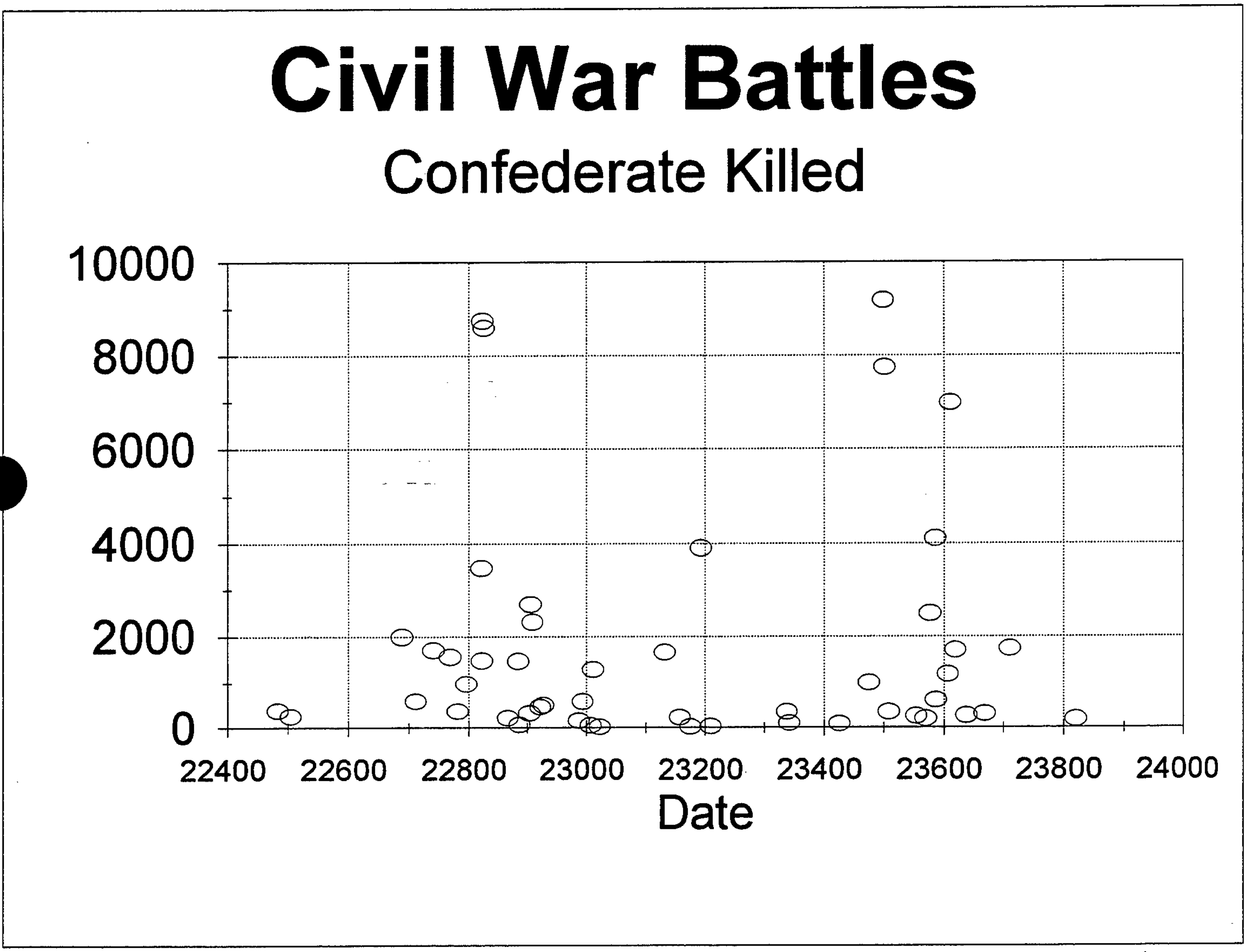


Figure XII.B.7

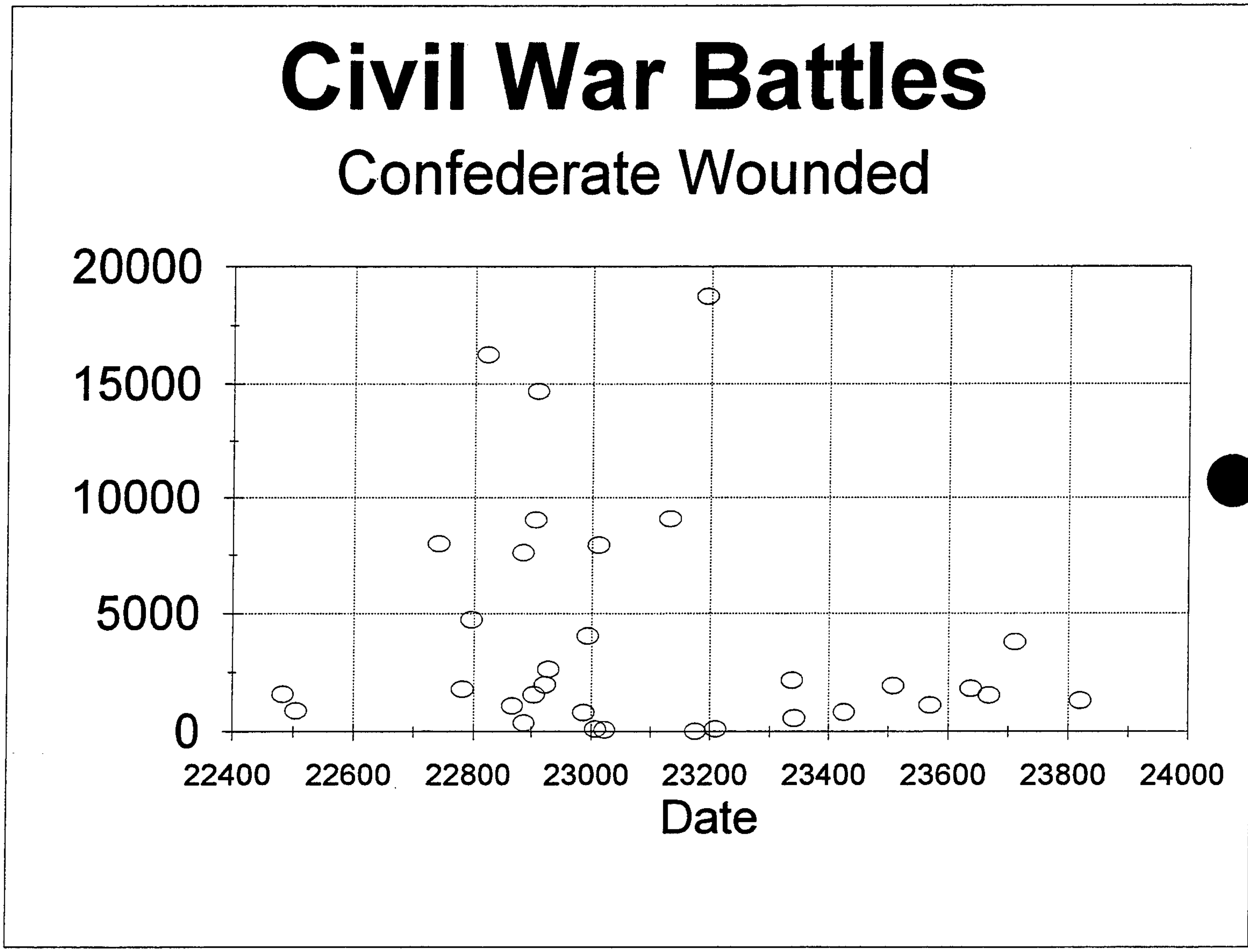


Figure XII.B.8

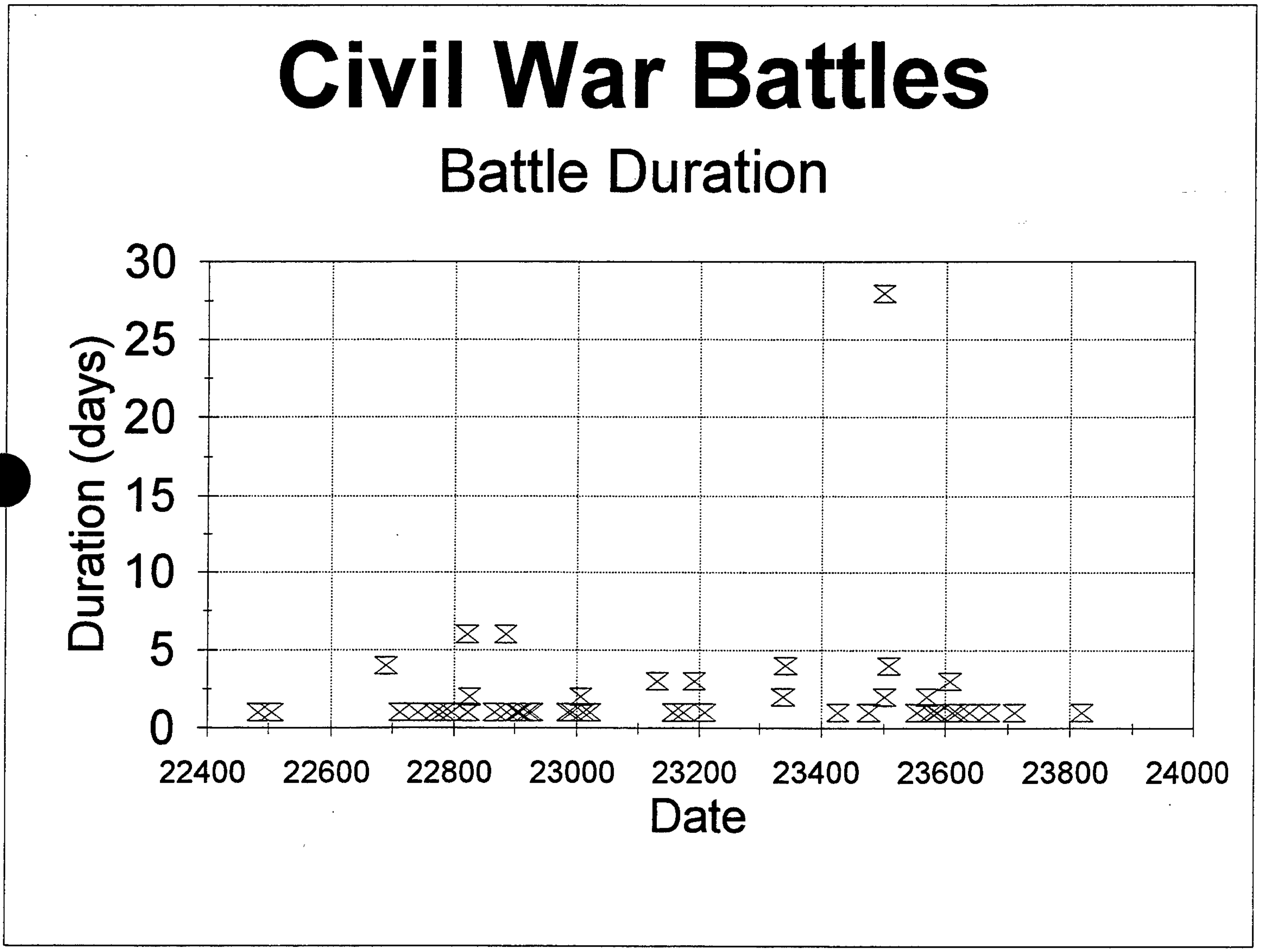


Figure XII.B.9

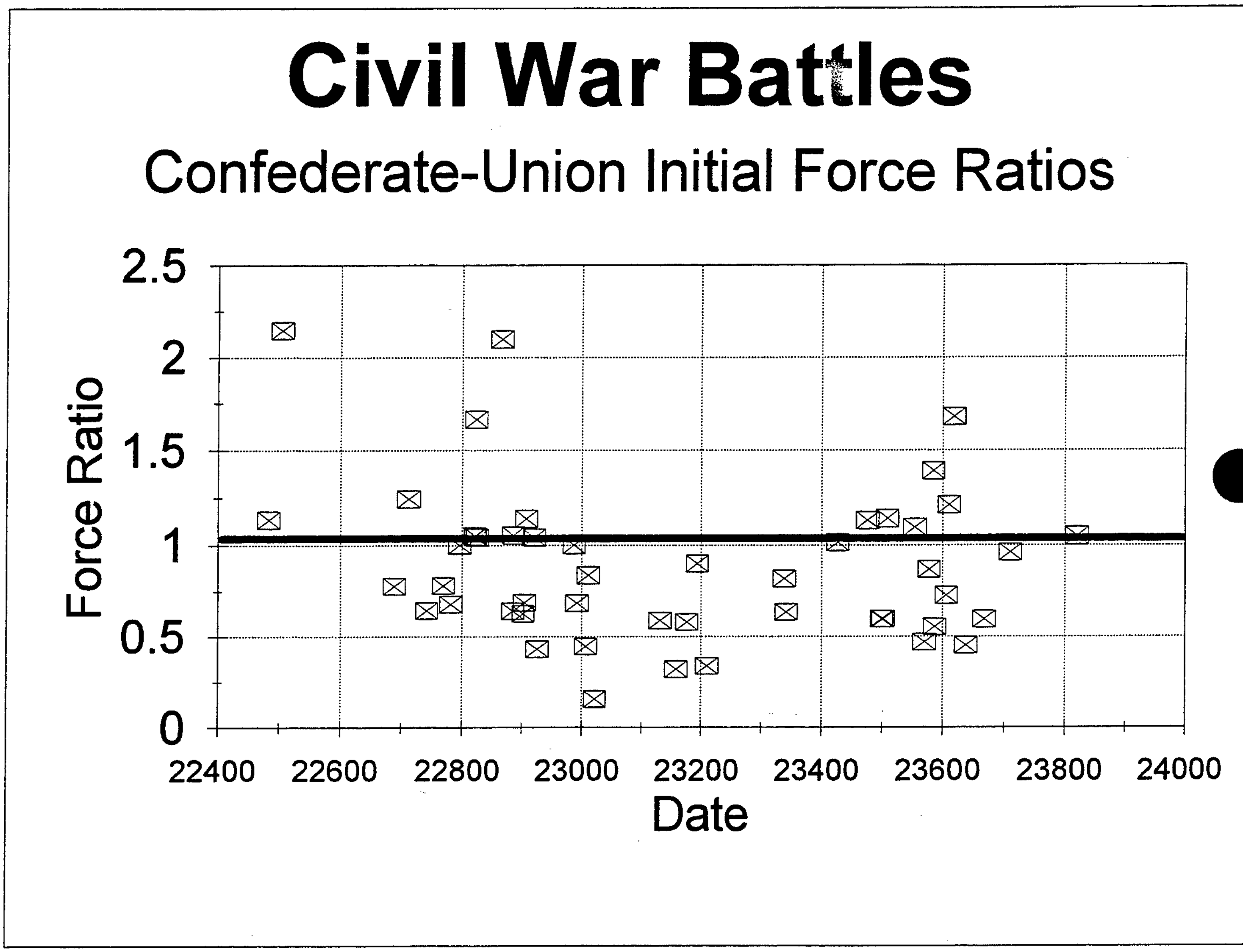


Figure XII.B.10

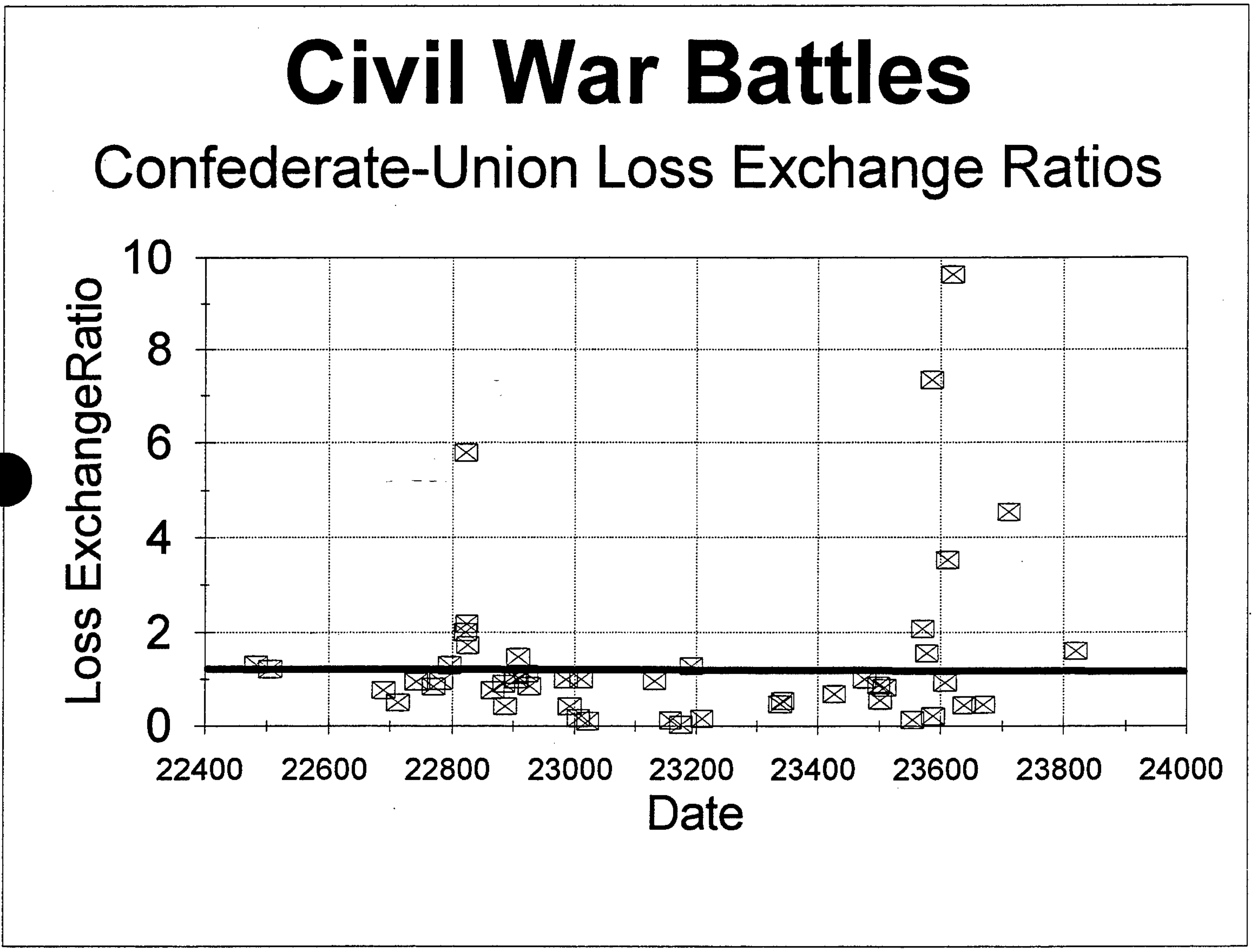


Figure XII.B.11

This does not however, paint a complete picture. An equally important quantity in this case, is the Fractional Exchange Ratio (F_{ER}) which shows the ratio of the relative losses for the two sides. If we define the fractional losses as

$$f_{l,U} \equiv \frac{\Delta U}{U_0}, \qquad \text{(XII.B-1)}$$

for the Union side, and

$$f_{l,C} \equiv \frac{\Delta C}{C_0}, \qquad \text{(XII.B-2)}$$

for the Confederate side, which show the fraction (on the interval [0,1],) of losses to each side, then the F_{ER} is just

$$F_{ER} \equiv \frac{f_{l,C}}{f_{l,U}}, \qquad \text{(XII.B-3)}$$

for the Confederate:Union ratio. We could equally well express the Fractional Exchange Ratio as the inverse of this (for the Union:Confederate ratio) if we so chose. Our interest here is in relative Confederate losses compared to relative Union losses so we define the F_{ER} this way. The student should note that this is an asymmetric representation because it accentuates large Confederate losses relative to Union losses. A symmetric representation would be the Logarithmic Fractional Exchange Ratio which is just the logarithm of equation (XII.B-3). We plot the F_{ER} versus date in Figure (XII.B.12). Note the greater number of high F_{ER}s in the latter part of the war, which, if we believe our data set to be representative, and we really have few other options if we are to try to draw any numerical insights, indicate a deterioration of tactical options or innovation to control losses. In all, we note F_{ER}s ≤ 1 in only about 45% of the cases, which clearly supports the theses of historians such as McWhiney and Jamieson. Further, no less than six battles have F_{ER}s greater than 2 and four of these are greater than 4! This number is significant compared to our total data set (49 battles) in demonstrating a Confederate willingness to accept high casualties despite their relative numerical inferiority and superiority as fighters indicated in the previous figure.

Before continuing, it is worthwhile to establish the relationship between fractional loss ratios, and F_{ER} and Lanchester Attrition theory. If we start with the general form of the state solution, equation (VII.B-5), slightly rewritten as

$$\alpha\left(C_0^n - C^n \right) = \beta\left(U_0^n - U^n \right). \qquad \text{(XII.B-4)}$$

If we expand C and U to first order in losses (ΔC, ΔU) and perform some minor algebra, this becomes

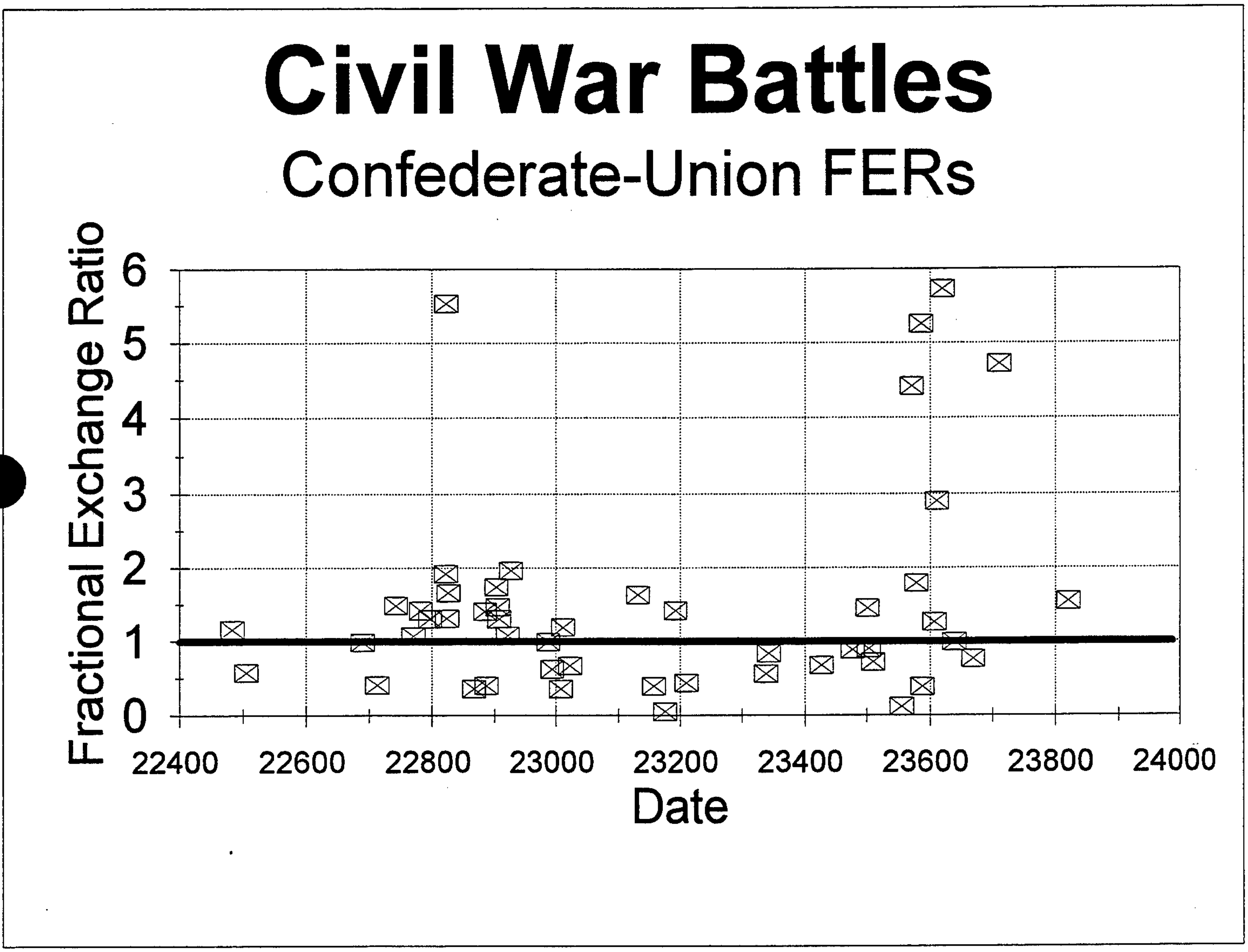


Figure XII.B.12

$$\alpha C_0^{n-1} \Delta C = \beta U_0^{n-1} \Delta U, \tag{XII.B-5}$$

which we may conveniently rewrite in the form,

$$\alpha C_0^n \frac{\Delta C}{C_0} = \beta U_0^n \frac{\Delta U}{U_0}. \tag{XII.B-6}$$

This equation may be recast immediately using equations (XII.B-1) and (XII.B-2) as

$$\alpha C_0^n f_{l,C} = \beta U_0^n f_{l,U}, \tag{XII.B-7}$$

which relates the fractional losses (as long as they are small,) to the initial force strengths and the attrition rates. Obviously, then the F_{ER} is just

$$F_{ER} = \frac{\beta U_0^n}{\alpha C_0^n}, \tag{XII.B-8}$$

again providing the losses are small. Note that the attrition order is preserved. It is interesting to note that Lanchester theory predicts that the F_{ER} of a battles should be a constant during its progress. We shall examine this issue in greater detail later.

Weiss also presents frequency distributions of force and casualty ratios. In keeping with his outline, we present the Confederate:Union initial force strength ratio frequency distribution in Figure (XII.B.13). The bin widths (0.2) for the frequency distributions are identical to those used by Weiss. Examination would lead us to speculate, except for the relative minima at force ratios on (0.8,1], that the distribution is Poisson or Gamma. Investigations of further distinctions, such as attacker/defender, or assault on fortified lines/other, might yield insights into the likelihood of attacking.

The equivalent distribution for C:U final force strength ratios is given in Figure (XII.B.14). This distribution has the same general form as the distribution of initial force strength ratios.

A similar consistency may be found in the frequency distributions of final to initial force strength ratios. These are given in Figures (XII.B.15) and (XII.B.16) for the Confederate and Union sides, respectively. In this case, the use of Weiss' bin sizes is ill chosen. (Weiss did not include these figures in his article.) Nonetheless, we shall investigate the relationship of final to initial force strengths in the next section.

We also include several other distributions for general interest. Figures (XII.B.17) and (XII.B.18) show the frequency distributions of Confederate (Union) final

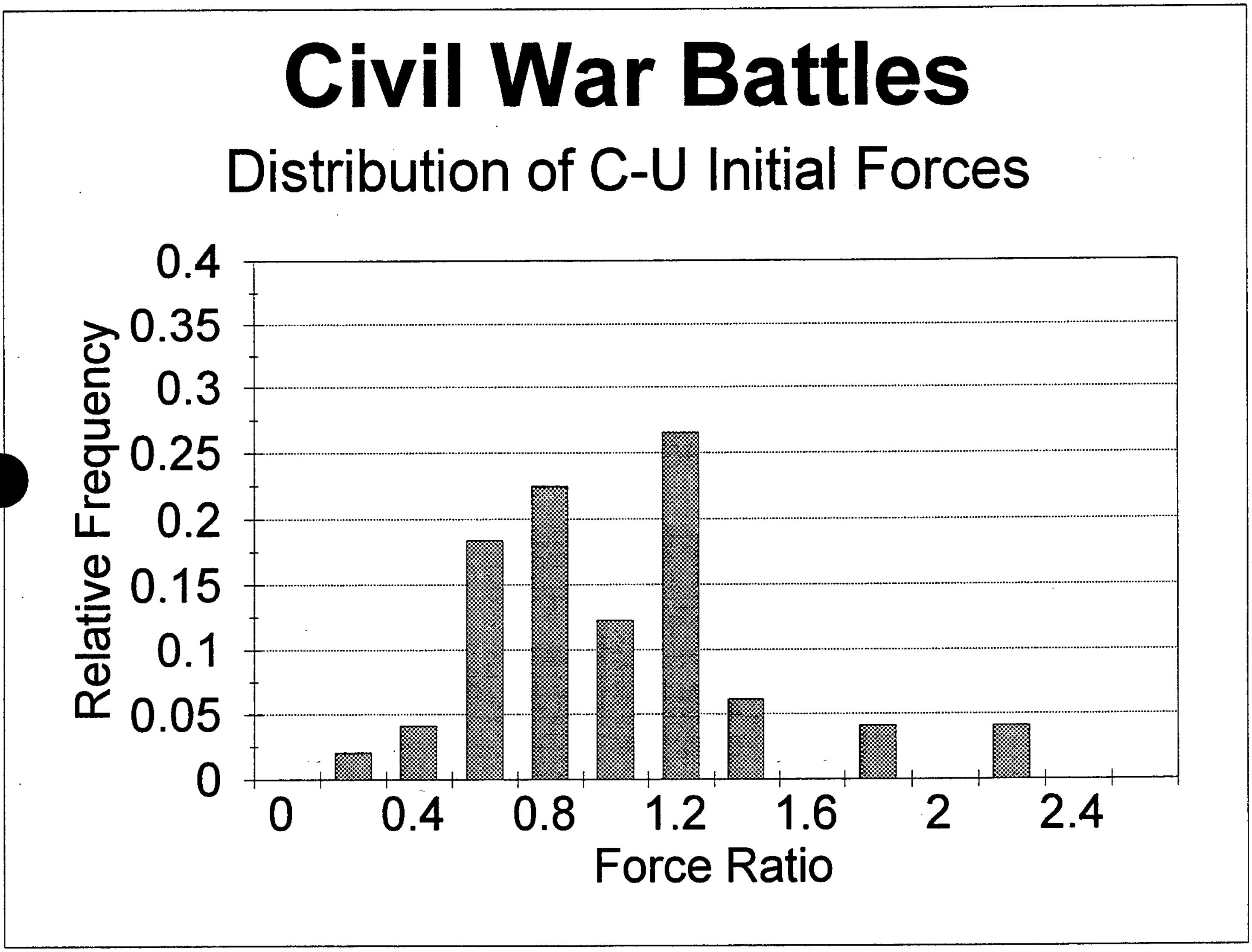


Figure XII.B.13

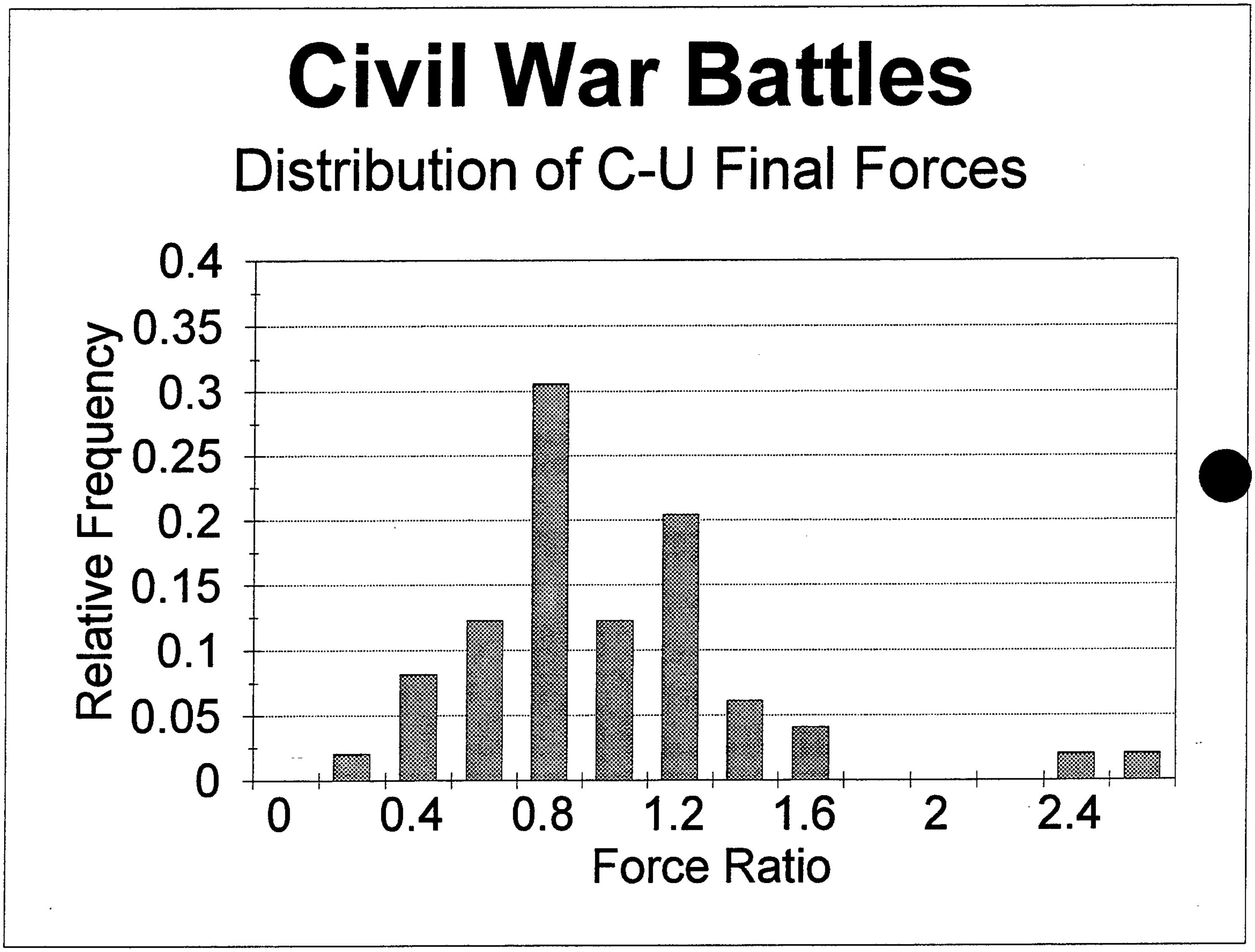


Figure XII.B.14

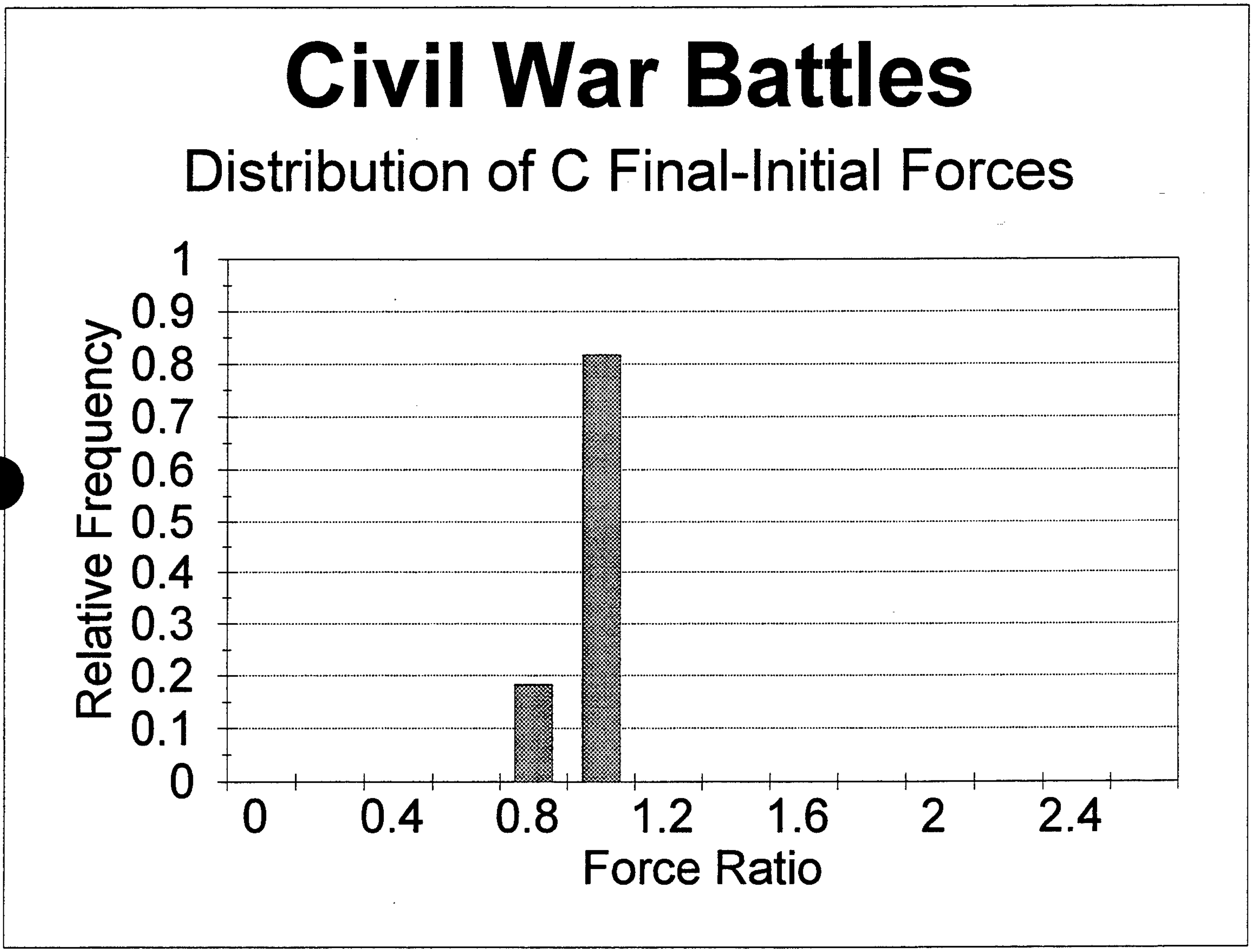


Figure XII.B.15

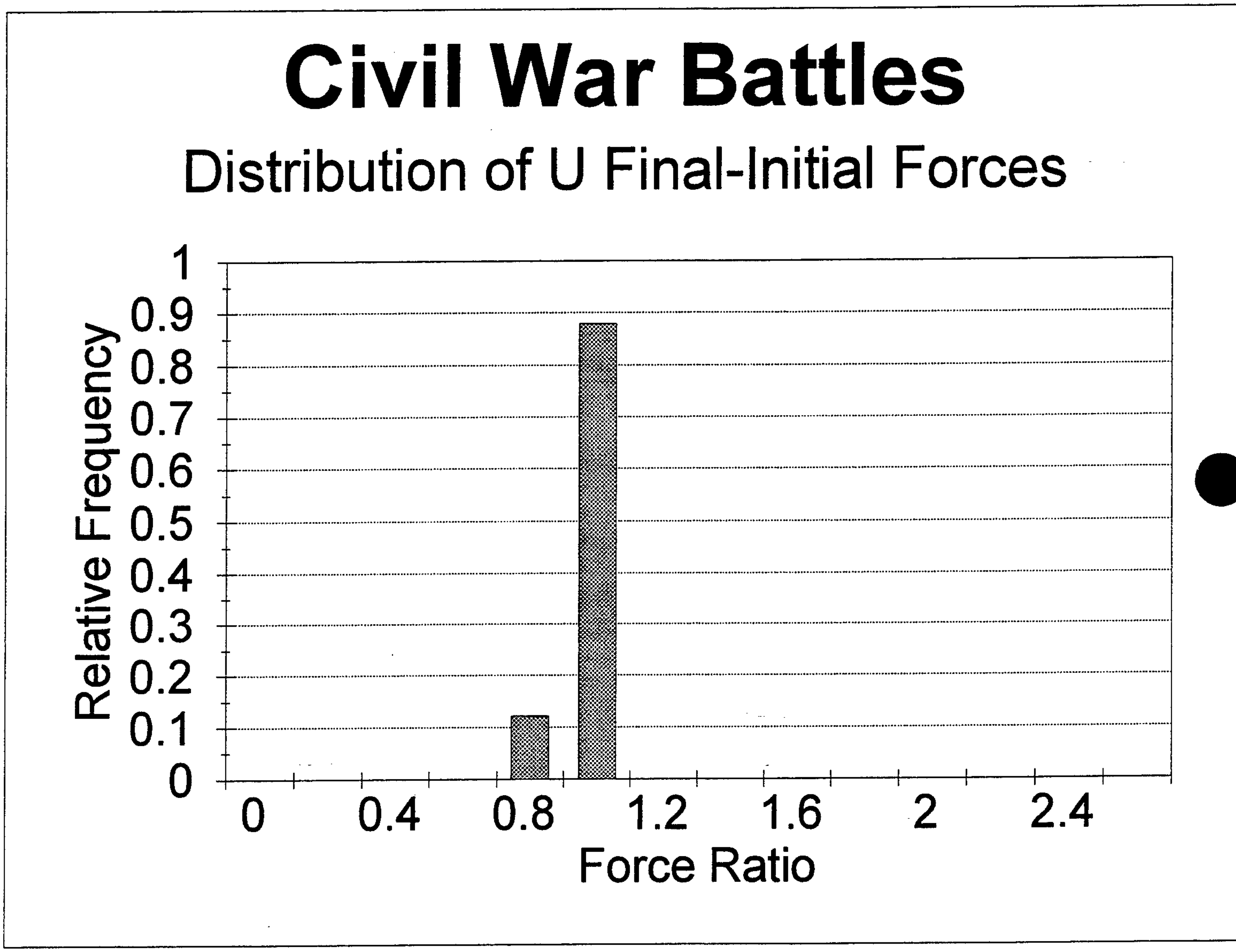


Figure XII.B.16

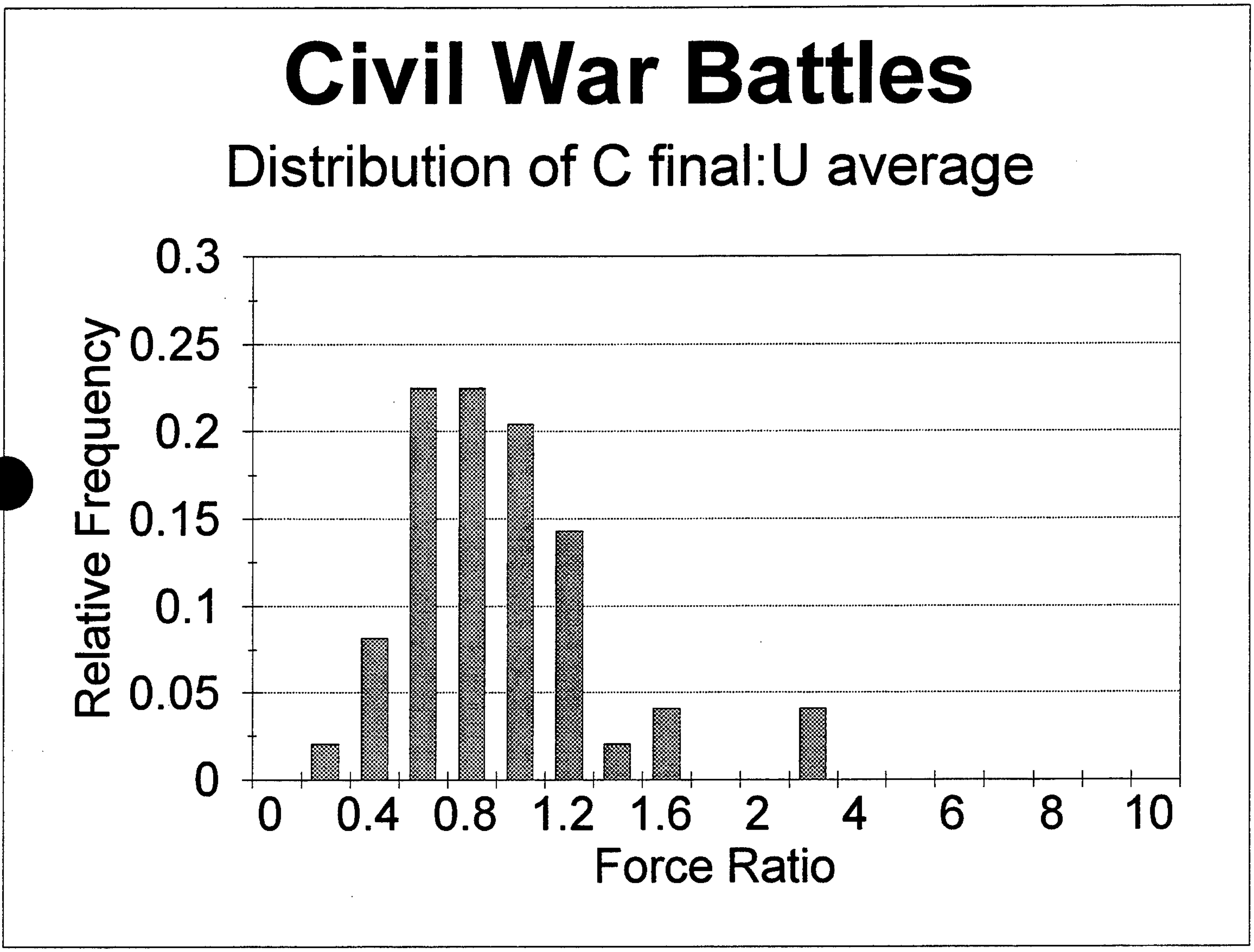


Figure XII.B.17

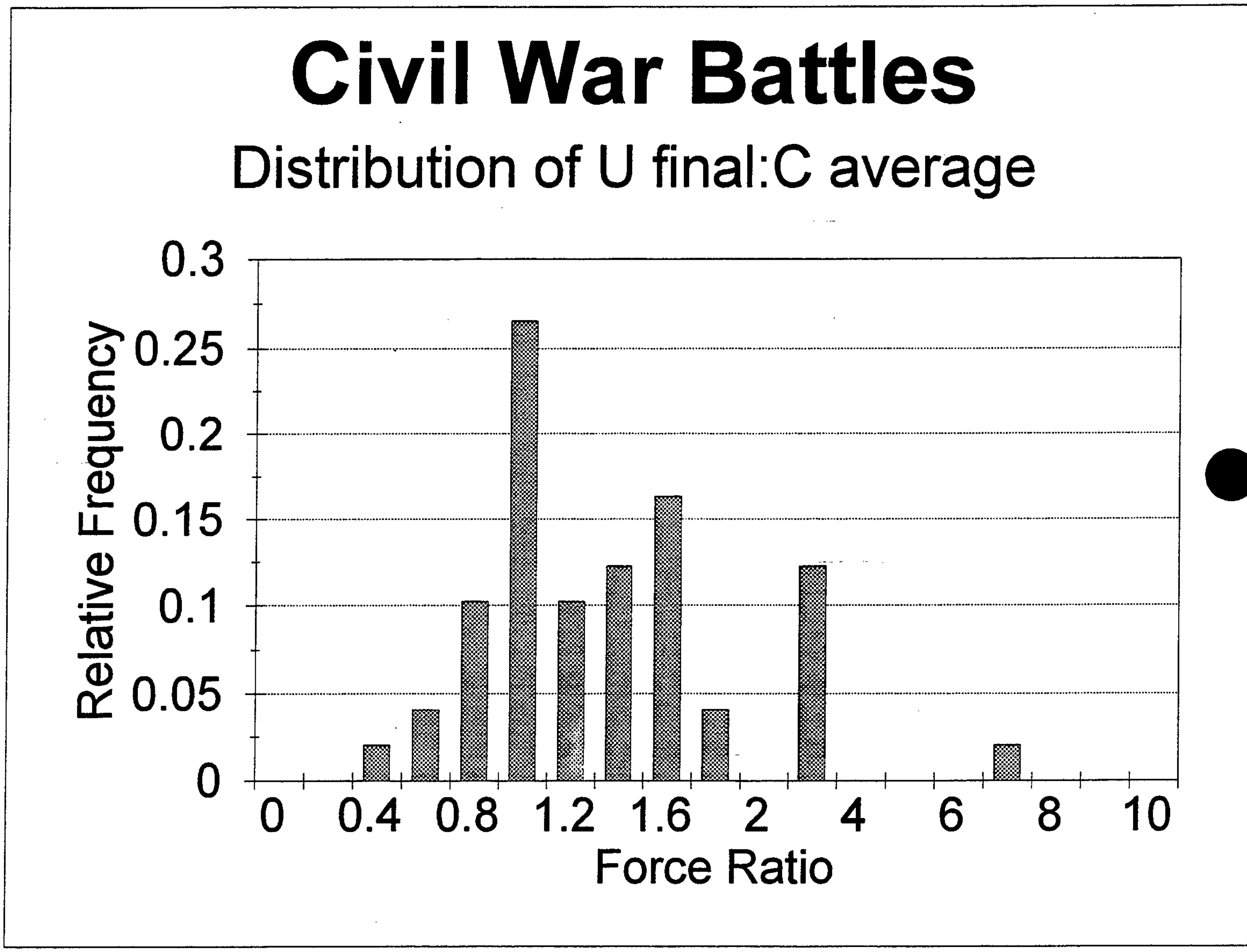


Figure XII.B.18

force strength to average Union (Confederate) force strength. The frequency distribution of the ratios of Confederate losses to Union losses is given in Figure (XII.B.19). None of these display any obvious type of pattern.

The frequency distribution of battle duration is given in Figure (XII.B.20). It would appear likely that this distribution is close to being negative exponential.

Next, in Figure (XII.B.21), we present the frequency distribution of C:U F_{ER}. The general shape again suggests a Poisson or Gamma distribution aside from the large frequencies in the 1.6, 5, and 6 bins.

Finally, we examine losses. In Figure (XII.B.22), we present a scatter plot of Confederate losses versus Union losses. This figure is a logarithmic plot, but is fairly clearly a symmetric pattern about a line with slope of approximately one. This behavior supports Weiss' findings about approximately equal losses on both sides.

In Figure (XII.B.23), we present a logarithmic scatter plot of C:U loss ratio versus U:C initial force ratio. We have not divided the data into two sets: attacks on fortified lines, and other; as Weiss did. Despite the scatter, there is a strong suggestion of a linear relationship with small negative slope. This is an interesting speculation. It implies that the Confederate forces were more effective against larger Union forces than against smaller ones. Does this further imply that Confederate leaders had a better command of Grand Tactics (Operational Art) than Union leaders had? This may be too strong an assertion, but it supports arguments of superior Southron generalship and makes a counter argument to McWhiney and Jamieson.

XII.C. Force Strengths and Attrition Order

In this section, we temporarily depart from our general outline of following Weiss' article to examine some of the behavior of force strengths and attrition order. As we have already noted in conjunction with Figures (XII.B.15) and (XII.B.16), the frequency distributions of final to initial force strength ratios are rather narrow. This merits additional consideration.

In Figure (XII.C.1), we present a scatter plot of Union final force strength versus Union initial force strength. As we noted in the preceding chapter, the degree of linearity shown in these Civil War data, indeed in all of our data sets, is striking. The equivalent scatter plot for the Confederate side, given in Figure (XII.C.2), is somewhat noisier, but similar and also striking. We may postulate that these data have the functional relationship,

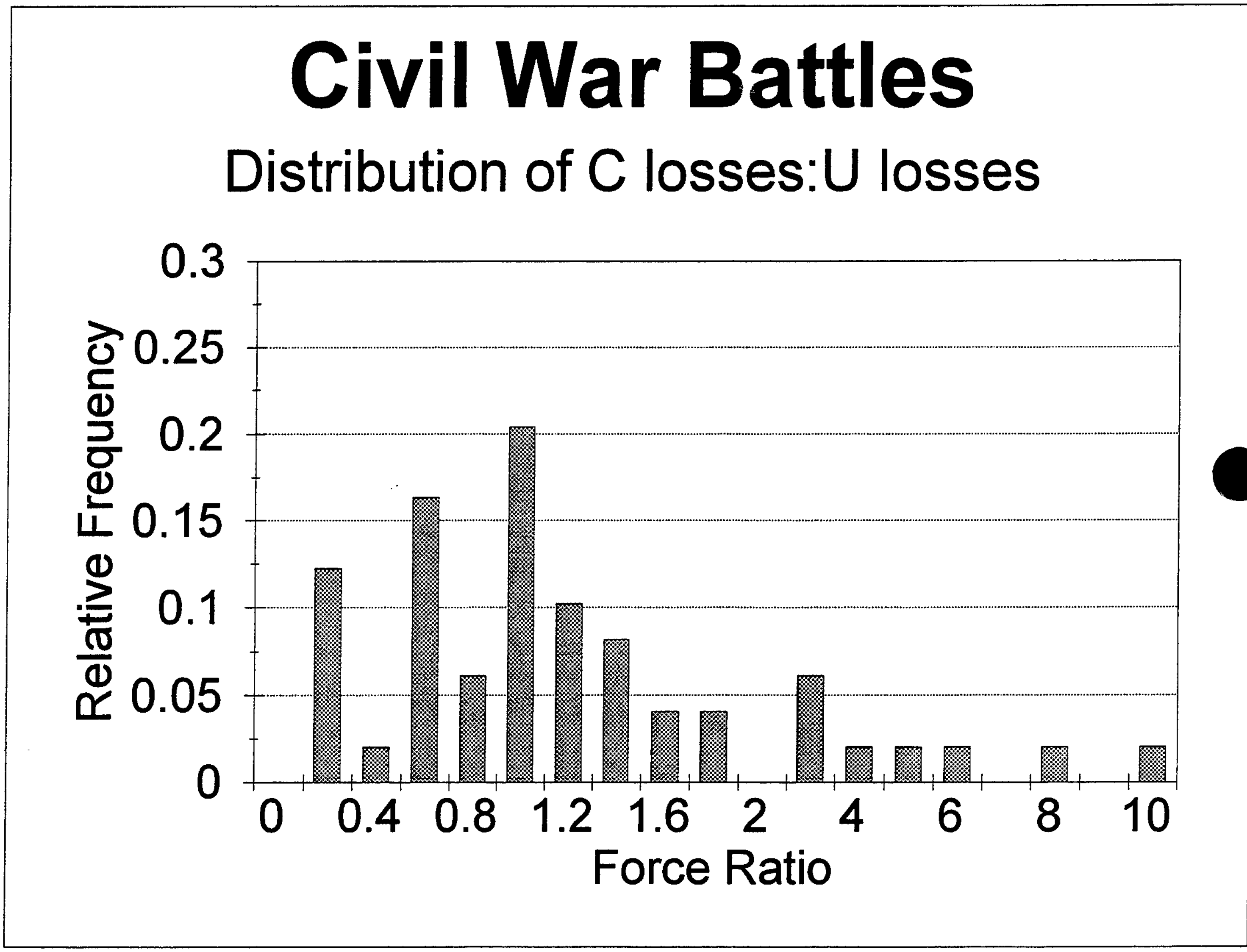


Figure XII.B.19

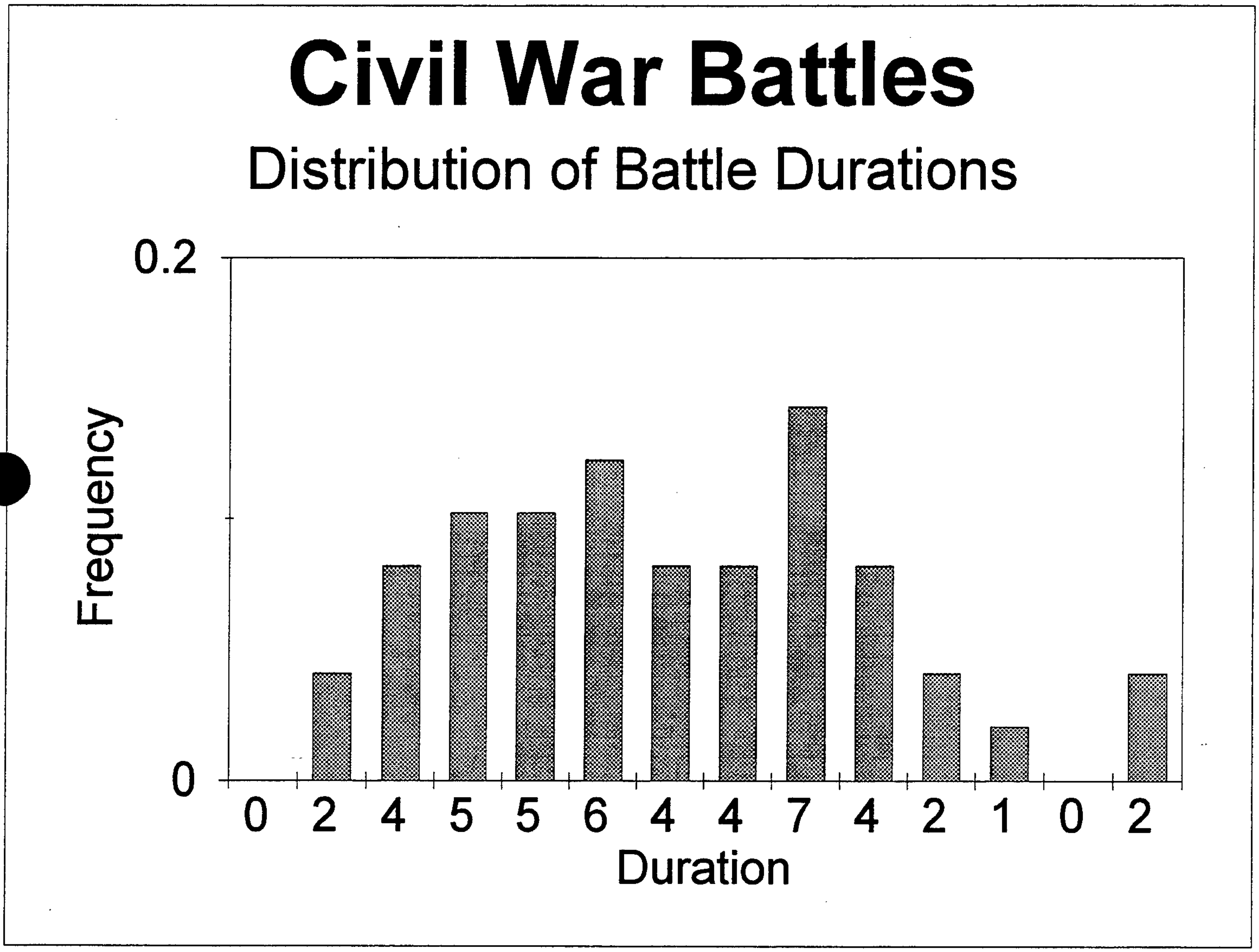


Figure XII.B.20

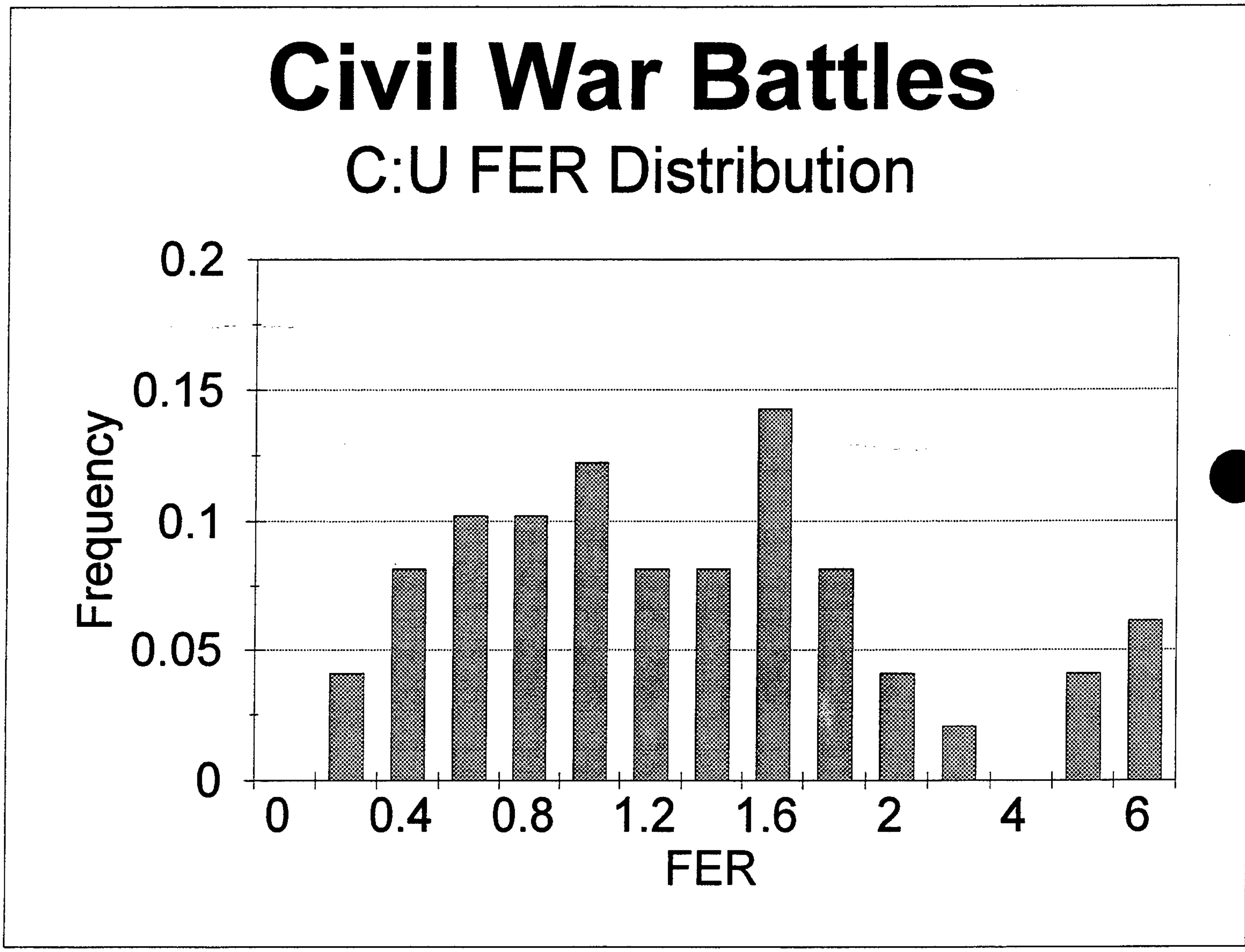


Figure XII.B.21

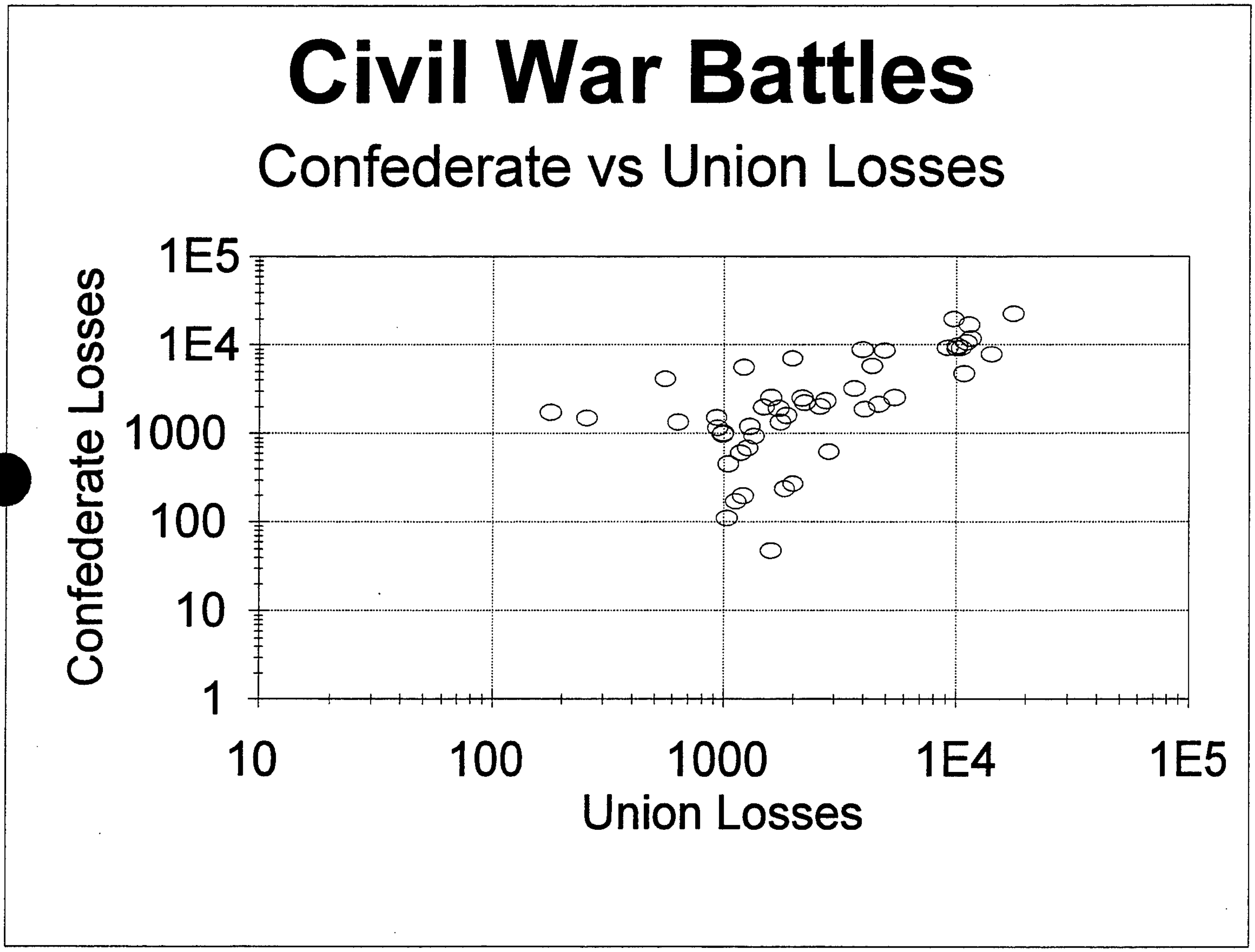


Figure XII.B.22

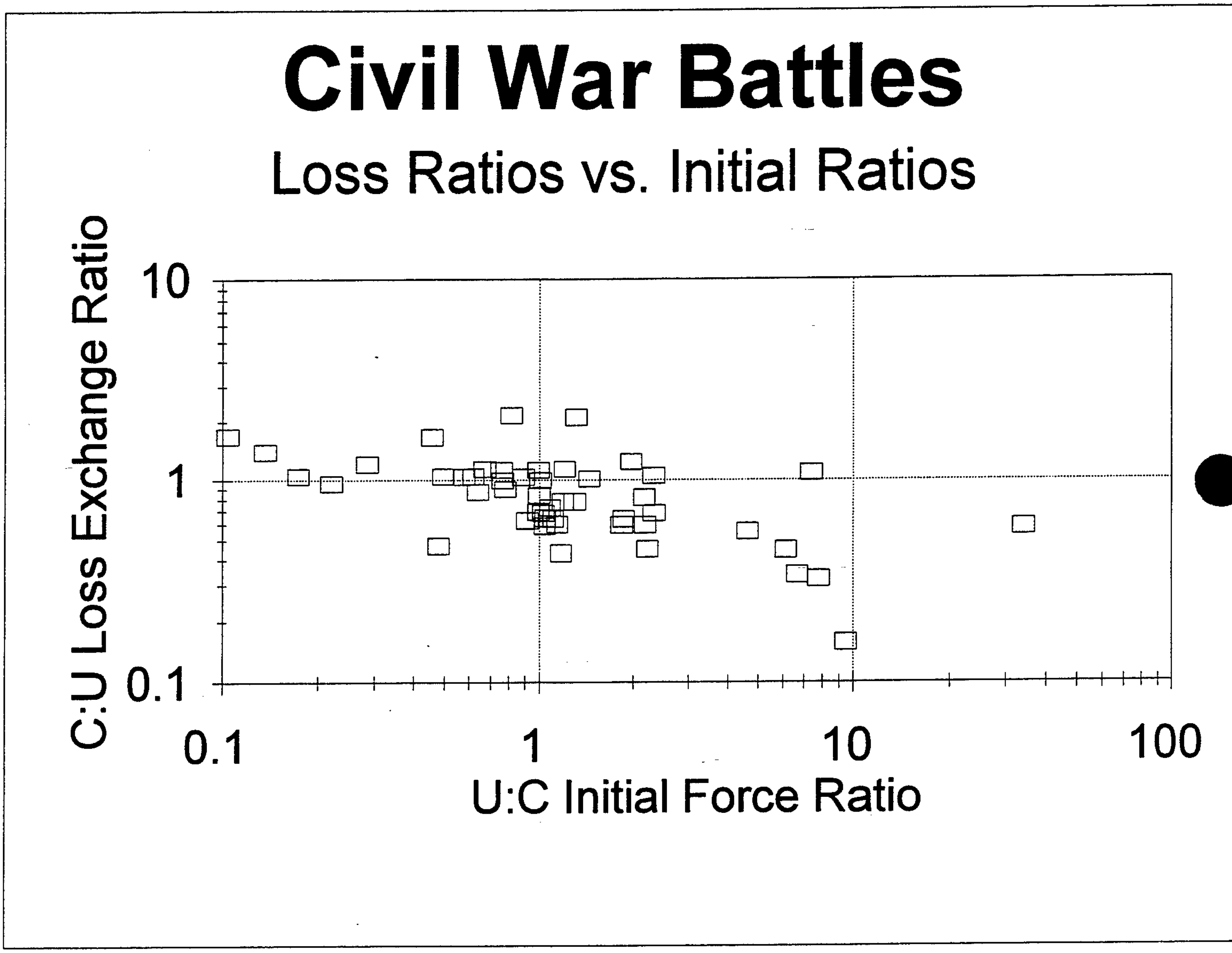


Figure XII.B.23

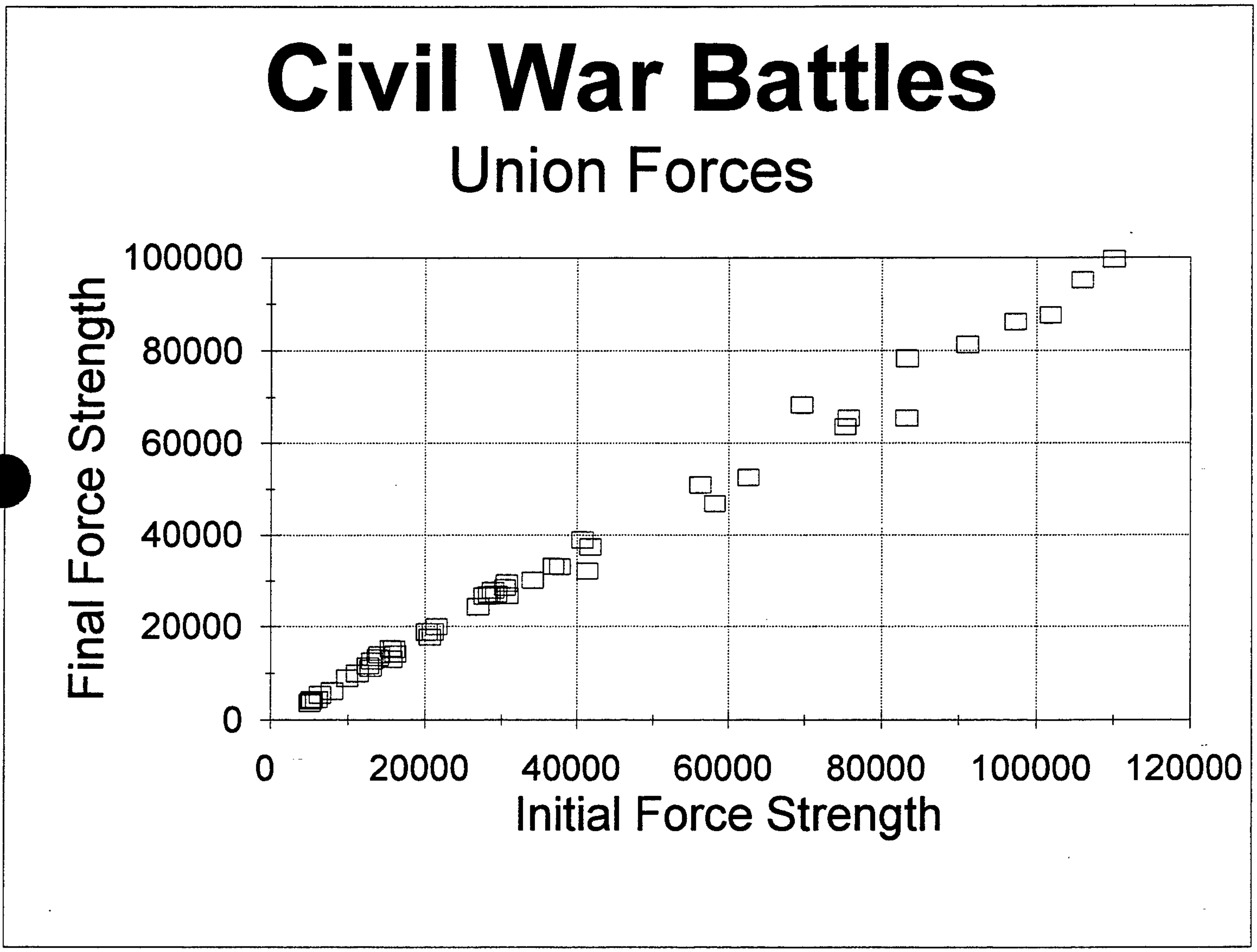


Figure XII.C.1

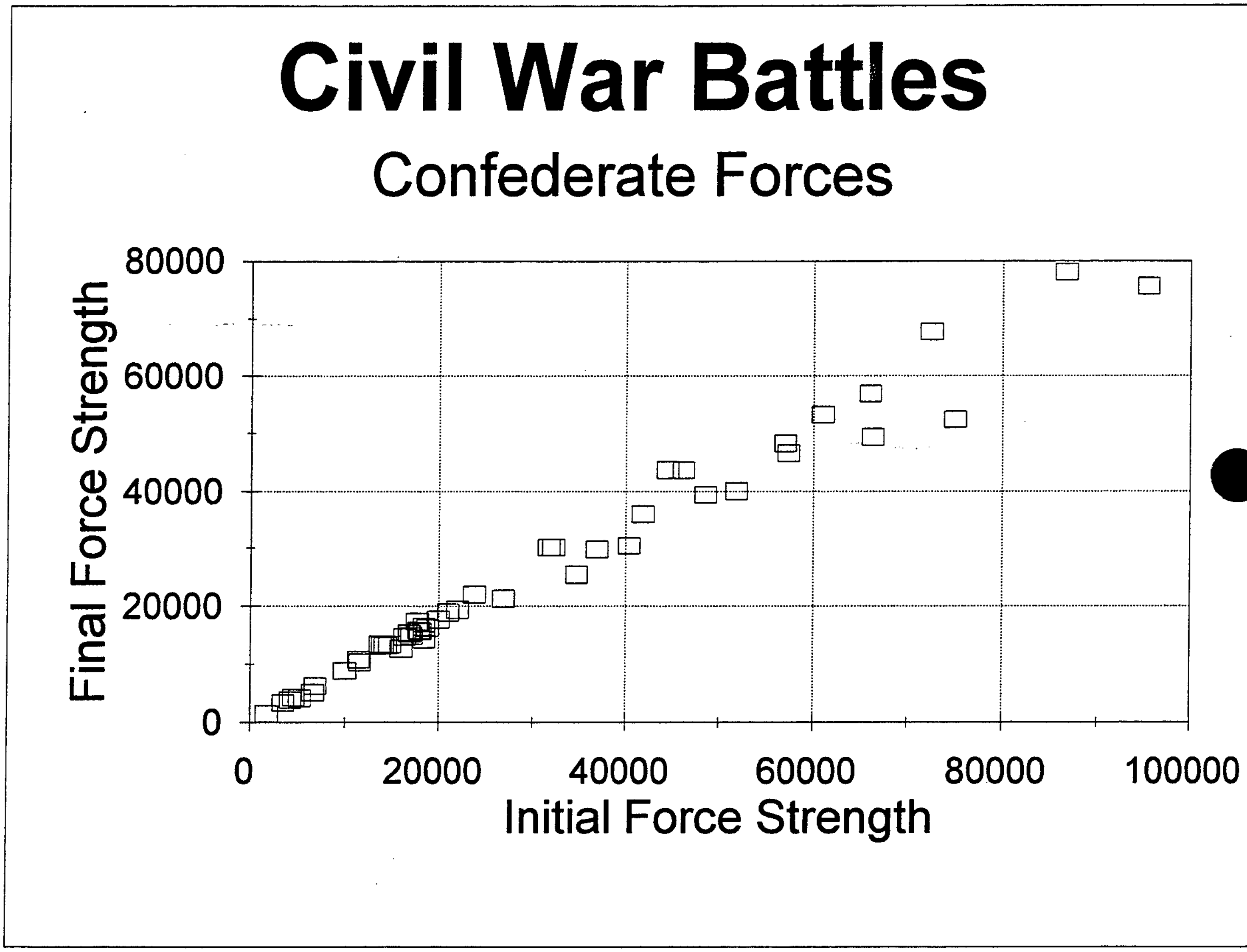


Figure XII.C.2

$$S_{final} = \sigma\, S_{initial}, \qquad \text{(XII.C-1)}$$

where S indicates force strength, and σ is the slope of the line. The line has no intercept value since we would expect there to be no losses for a force of zero initial strength. If we curve fit these data, we obtain the slopes given in Table (XII.C.1) We

Table XII.C.1 Homogeneous Final:Initial Force Strength Linear Relationships

Side	Slope	Slope Standard Error	R^2
Union	0.8852	0.0069	0.9926
Confederate	0.8405	0.0116	0.9766

may glean several insights from this information. First, on the average, Union forces lost about 11% of their initial strength per battle while Confederate forces lost about 16% of theirs per battle. This is another piece of evidence supporting the arguments advanced by McWhiney and Jamieson. The small slope standard errors and the large R^2 values[f] indicate the relative goodness of the fit and strongly supports the concept of functional relationship between initial and final force strengths implicit in Lanchester Attrition Theory. The relatively large standard error and smaller R^2 for the Confederate side reflects the greater spread in the data as shown in Figure (XII.C.2) as compared to Figure (XII.C.1).

We may also examine the behavior of Confederate (Union) final force strength versus Union (Confederate) initial force strength. These scatter plots are given in Figures (XII.C.3) and (XII.C.4). We can immediately see that there is considerably greater scatter in these cross force plots than in Figures (XII.C.1) and (XII.C.2). If we curve fit these data using Equation (XII,C-1), we obtain the results given in Table (XII.C.2). The values of the slopes clearly reflect that C:U initial force ratios were generally less than one (57%), and that Confederate losses were generally relatively larger than Union losses. The relatively smaller values of R^2 as compared to those in Table (XII.C.1) indicate the lesser tightness of the relationship to the data, but their essential equality hints at correlation.

---

[f] Recall that R^2 is defined on the interval (0,1) and the closer its value to one, the better the correlation of data and fit.

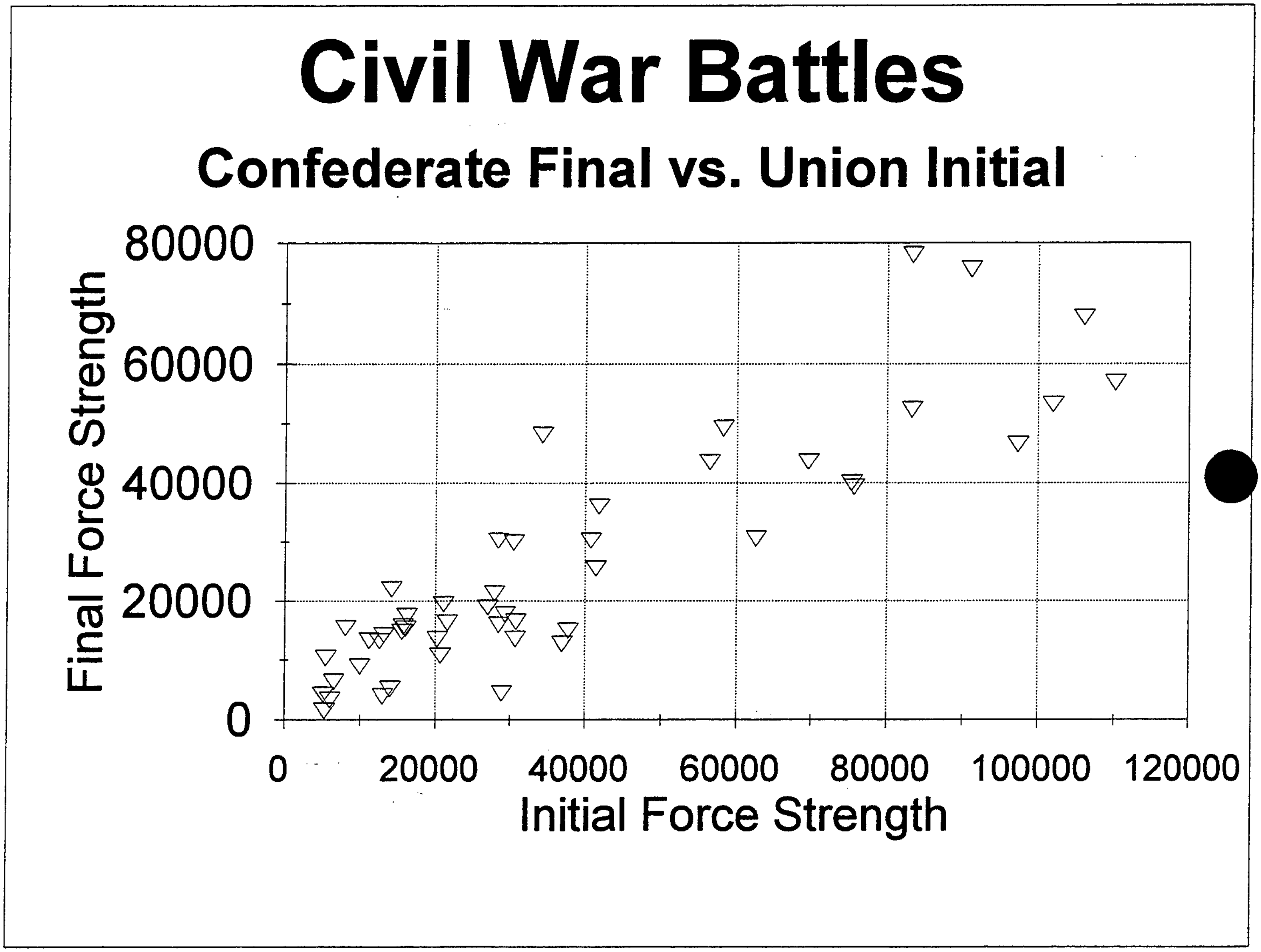


Figure XII.C.3

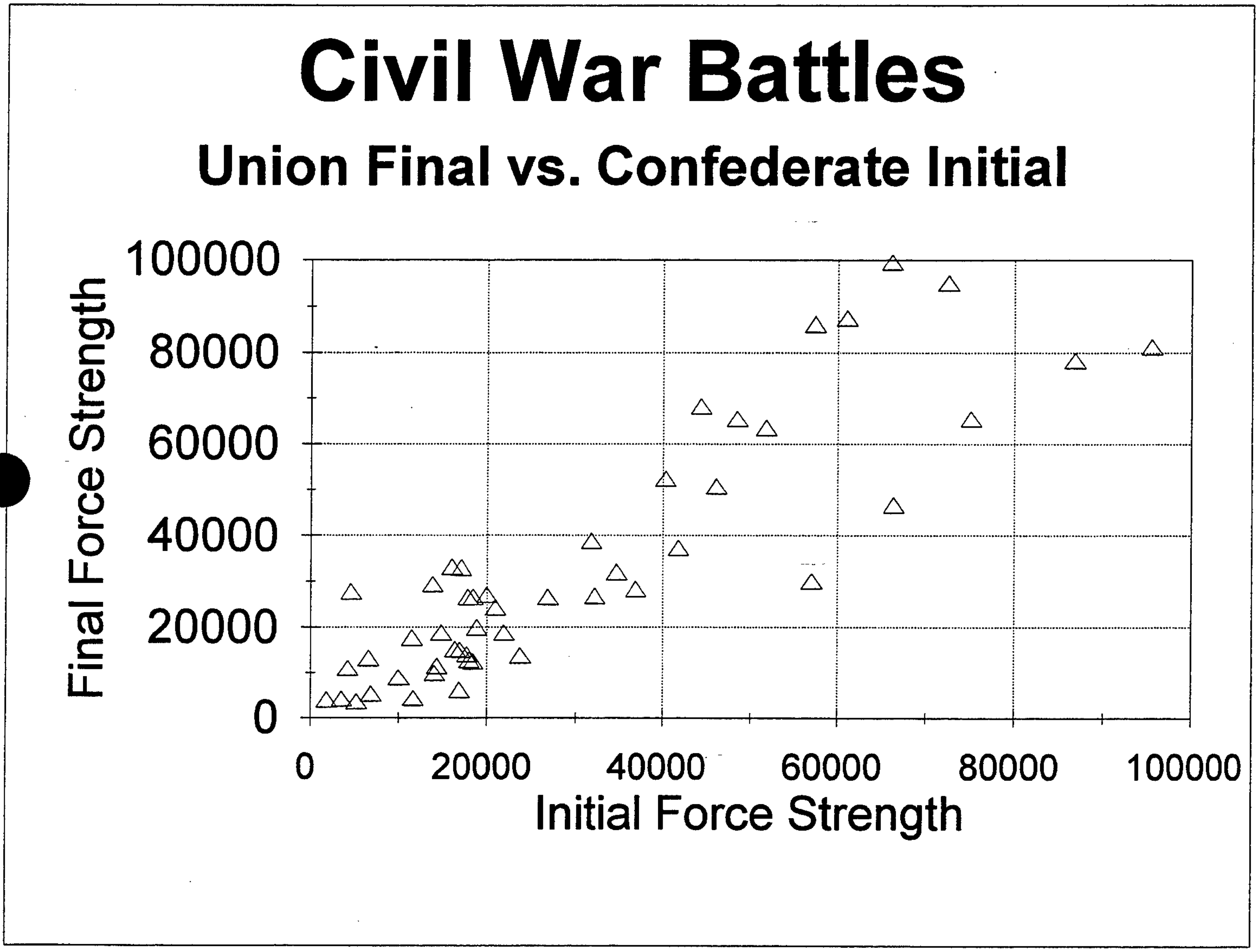


Figure XII.C.4

Table XII.C.2 Heterogeneous Final:Initial Force Strength Linear Relationships

Initial	Final	Slope	Slope Standard Error	R^2
Union	Confederate	1.0684	0.0476	0.7741
Confederate	Union	0.6415	0.0275	0.7735

We have previously examined the attrition order of these data with our approximately integrated differential equation (AIDE), which for our problem takes the form,

$$\ln\left(\frac{2\Delta U}{U_0 C_0 + UC}\right) = \ln(\alpha\tau) + (1 - n)\ln(U_0), \qquad \text{(XII.C-2)}$$

for the Union, and

$$\ln\left(\frac{2\Delta C}{U_0 C_0 + UC}\right) = \ln(\beta\tau) + (1 - n)\ln(C_0), \qquad \text{(XII.C-3)}$$

Table XII.C.3, Attrition Orders and Average Attrition Rates

Side	Attrition Order	Order Standard Deviation	Average Attrition Rate	Rate Standard Error	R^2
Total	2.097	0.105	0.179	1.088	0.444
Union	1.949	0.146	0.077	0.879	0.475
Confederate	1.865	0.082	0.176	0.497	0.699

for the Confederacy. The total is calculated by combining the two respective data sets. Now, since we fully recognize that there may be differences between Union and Confederate tactics, it is useful to separately examine the attrition order (and average attrition rate) of each side. We present these in Table (XII.C.3), and plots of the basic

data (left hand sides of equations (XII.C-2) and (XII.C-3) versus initial force strength) in figures (XII.C.5)-(XII.C.7). We have added a line which corresponds to an attrition order of 2 for illustrative purposes only.

If we examine these attrition orders, it seems reasonable to postulate that overall, for both sides, an attrition order of two (i.e. Quadratic Lanchester attrition,) is within a standard deviation of being accurate. Examination of the individual sides reveals smaller attrition orders, but for the Union, an attrition order of two is within a standard deviation. This is not the case for the Confederacy. Alternately, if we view our available choices of attrition order as 1, 3/2, or 2, then clearly these battles and engagements can be viewed as having attrition orders of 2. The equivalent plot for Willard's equation is given in figure (XII.C.8), for comparison. If we add the logarithm of the initial force strength to equations (XII.C-2) and (XII.C-3), and multiply by minus one, the result is a linear equation where the slope is the attrition order. We replot the Union and Confederate data using this adjusted AIDE in figures (XII.C.9) and (XII.C.10). The $n = 2$ illustrative line has also been plotted.

Lacking a rigorous theory to explain attrition orders other than these, we can accept from these plots that the battles and engagements in our Civil War data set may be approximately described using the Quadratic Lanchester equations. The pundit may claim that the R^2 values are fairly low, and this is indeed the case. We may reply however, that the R^2 for this data set using Willard's equation is smaller yet (0.3466.) The pundit may also object that for this data set, we also have battle duration, so that we may explicitly remove the τ from the slope and calculate the average attrition rate. If we do this, we find greater error, so we are forced to conclude that the quantity attrition rate times duration is more representative than are the two separately.

The values of the attrition rates are surprising. It appears that the rate at which Union forces could inflict losses on Confederate forces is, on the average, more than twice as large as the rate that Confederate forces could inflict losses on Union forces. The standard errors are quite large however, so the matter demands closer attention. If we accept that the attrition order of these battles and engagements is approximately two, then we can calculate the individual attrition rates using the approximately integrated differential equations for attrition order (n) of two. We present the scatter plot of these data in figure (XII.C-11). There is no obvious pattern here, but we can observe that the Confederacy's attrition rate exceeded 0.2 in ten cases, while the Union's only exceeded this value in eight cases. In only two cases did both exceed 0.2 in the same battle. An interesting further investigation would be to examine the difference between attacker and defender. If we examine the relationship between attrition rate, and initial force strengths, shown in figure (XII.C-12), we note the interesting trend that the larger the initial force strength, the less likely that the attrition rate would be large.

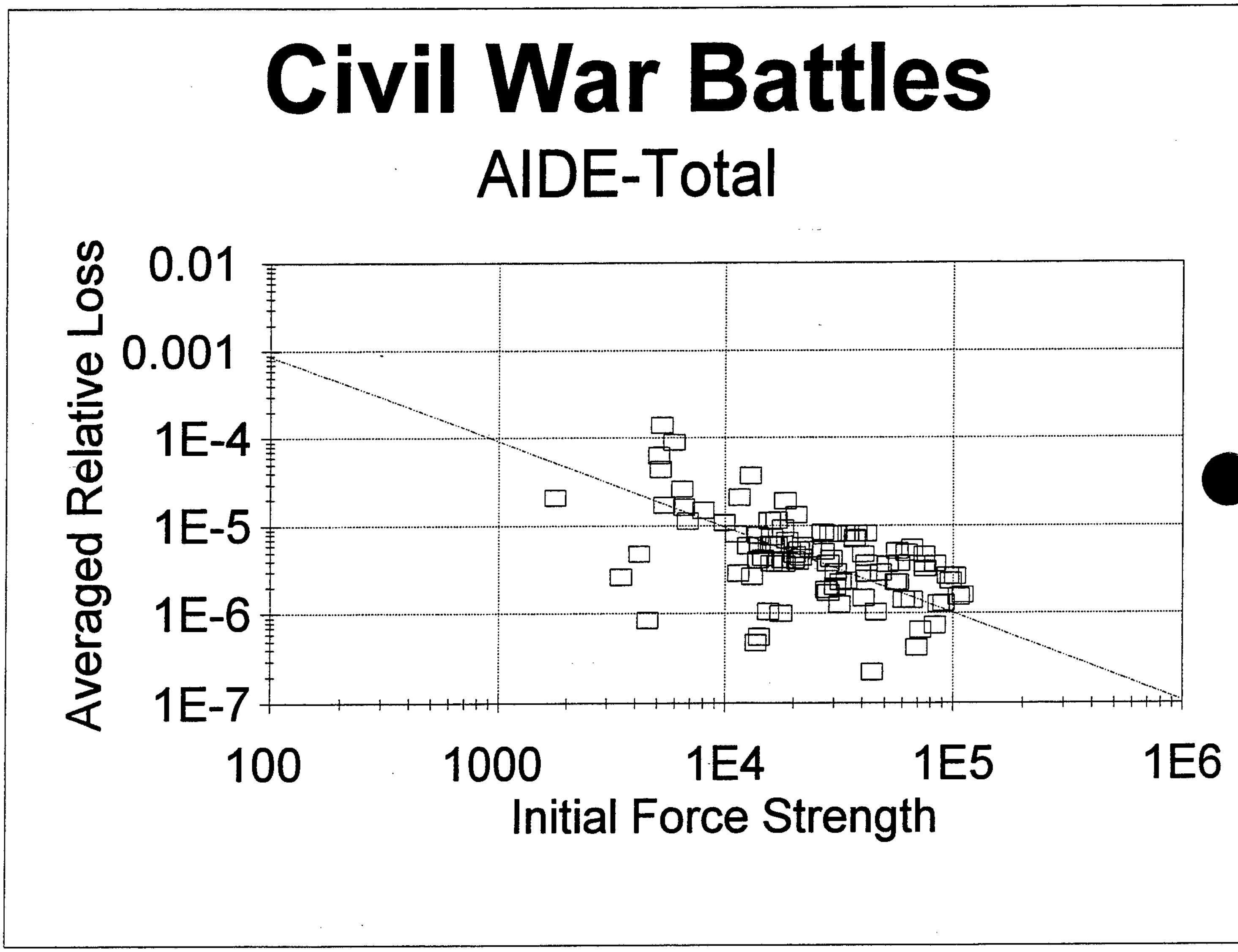


Figure XII.C.5

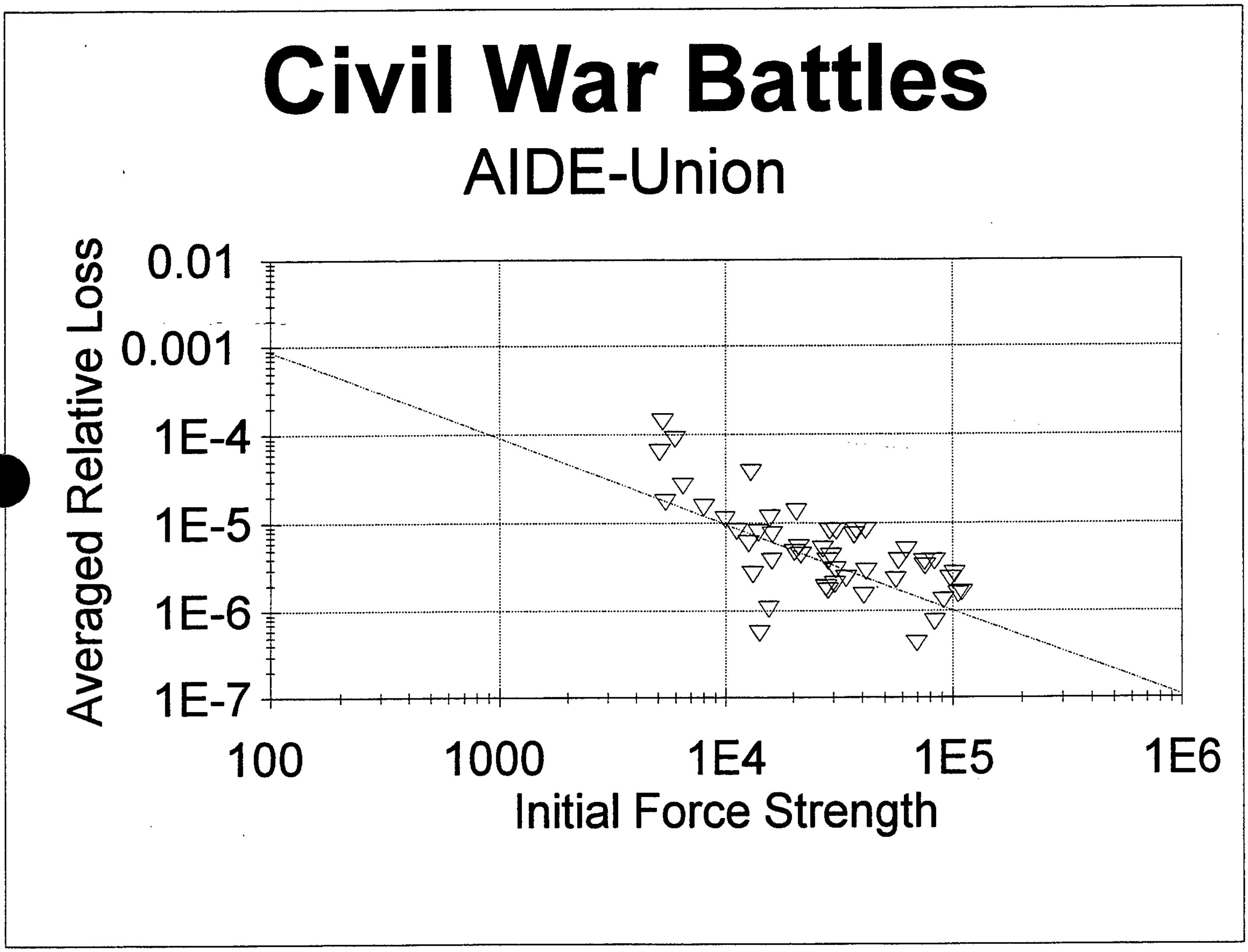


Figure XII.C.6

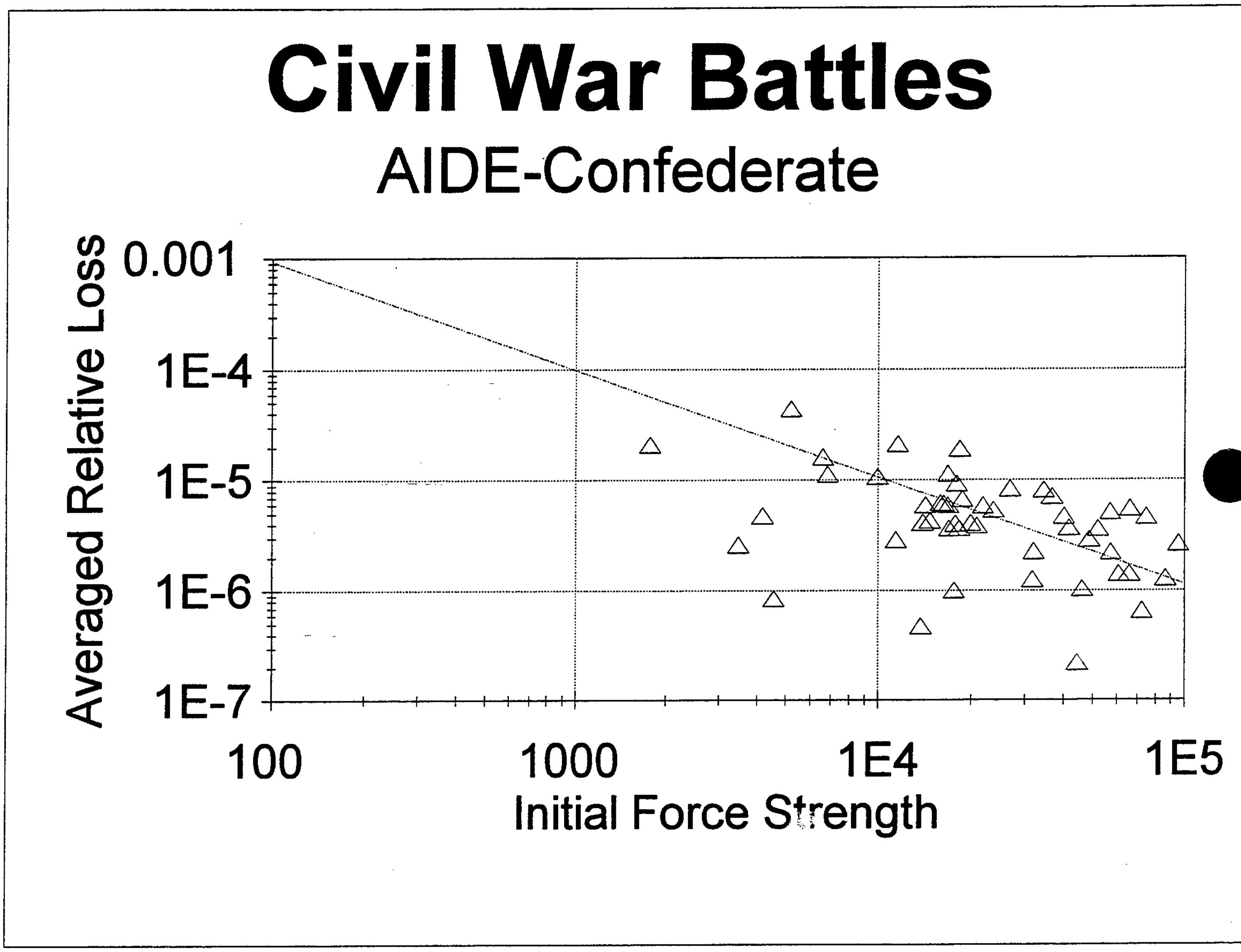


Figure XII.C.7

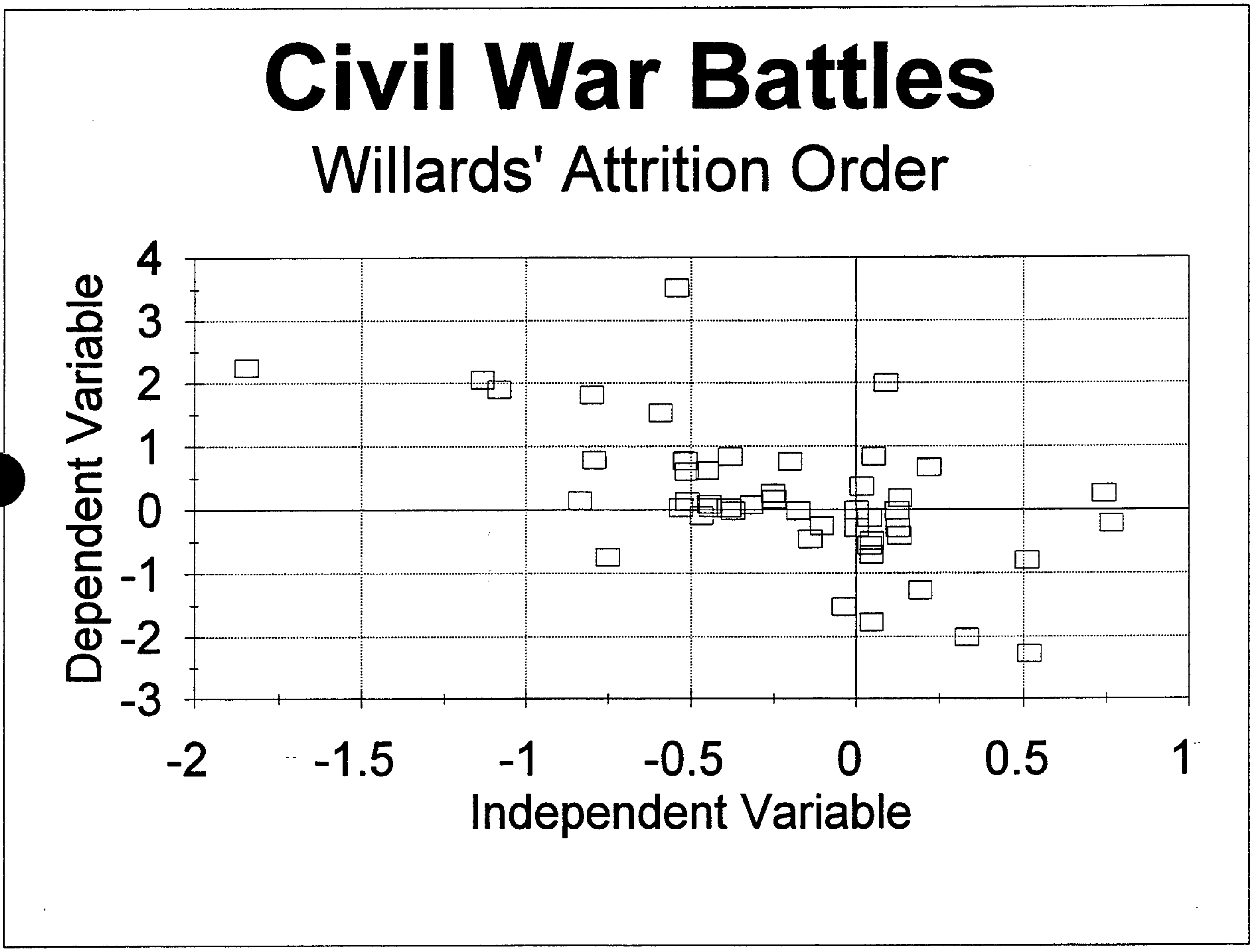


Figure XII.C.8

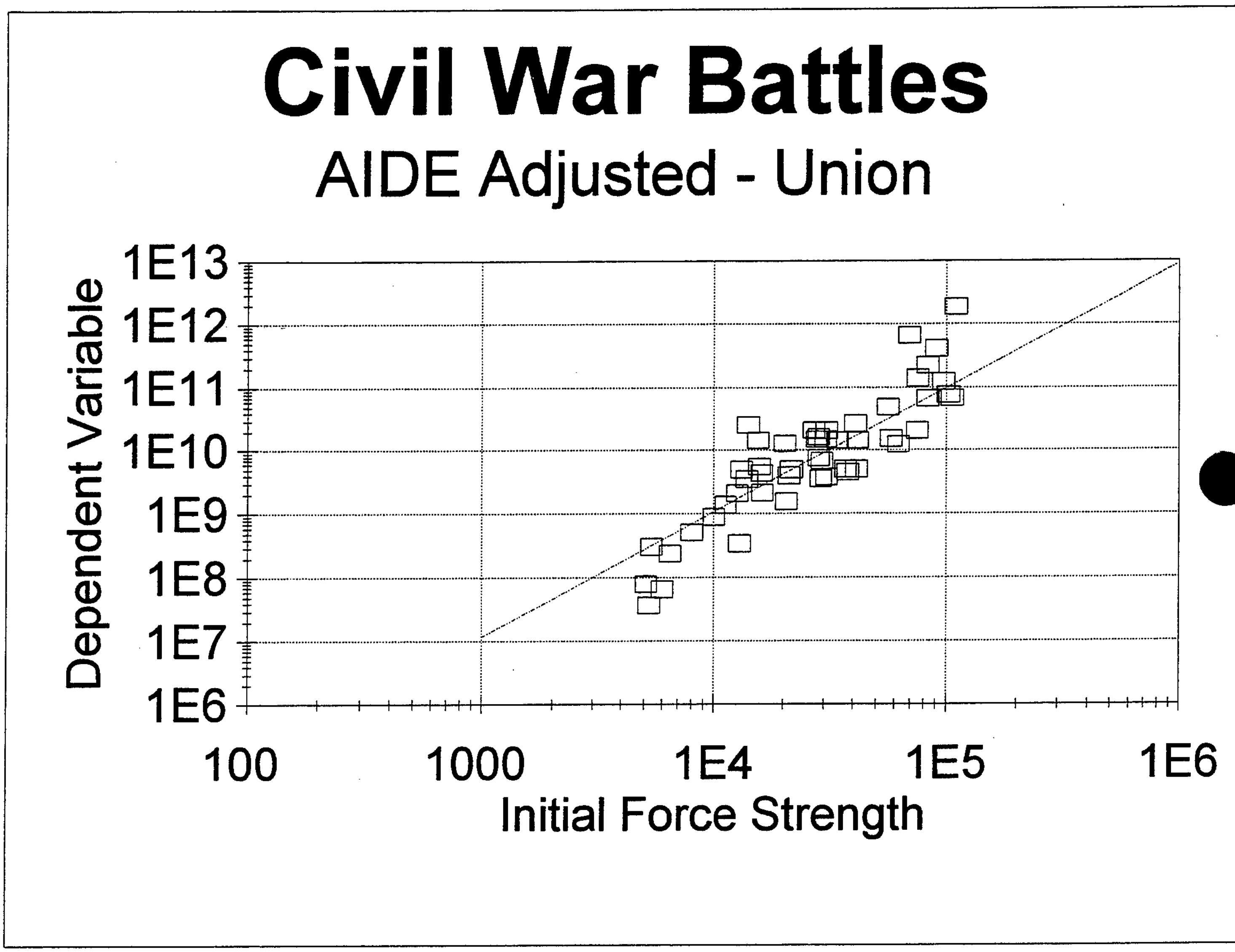


Figure XII.C.9

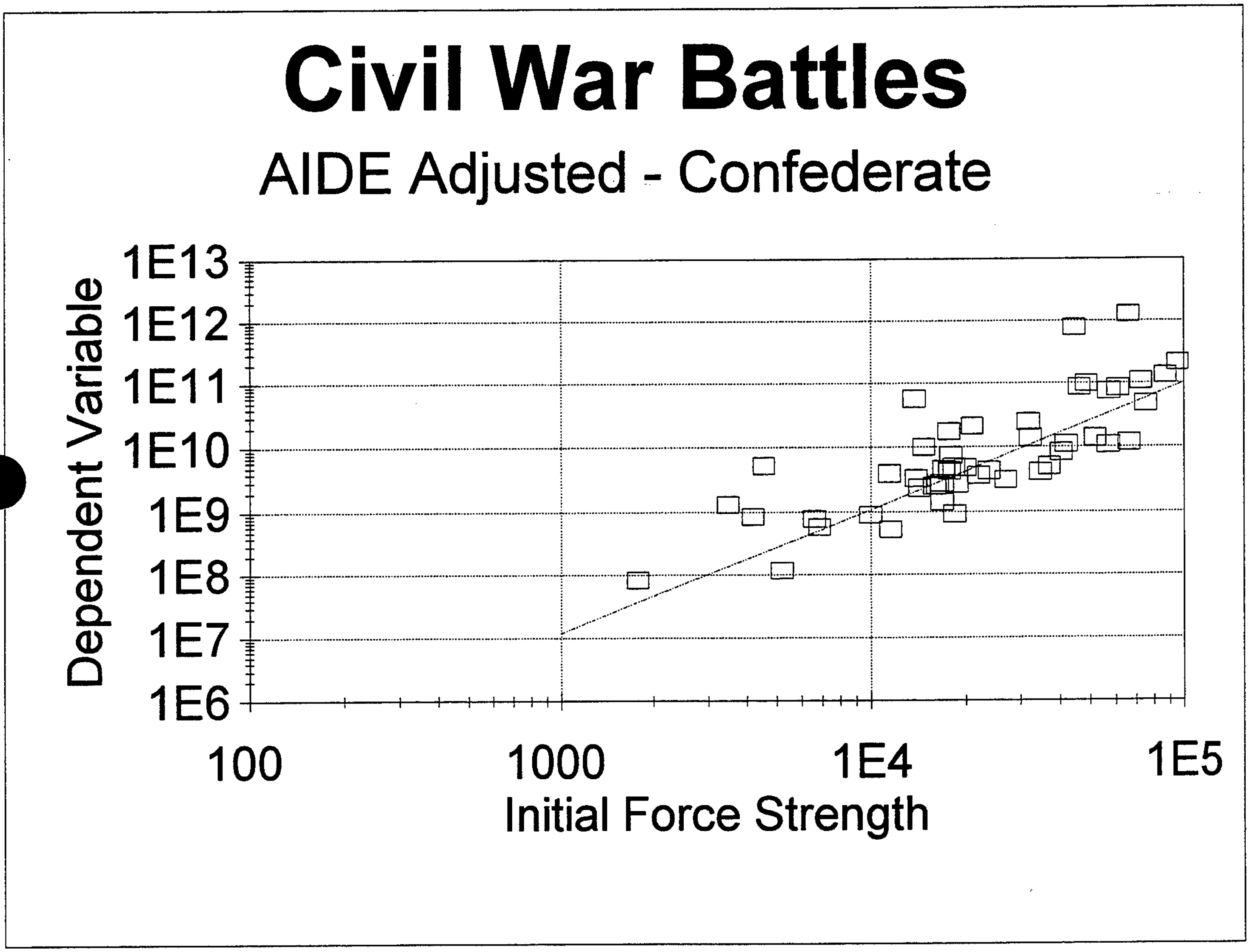


Figure XII.C.10

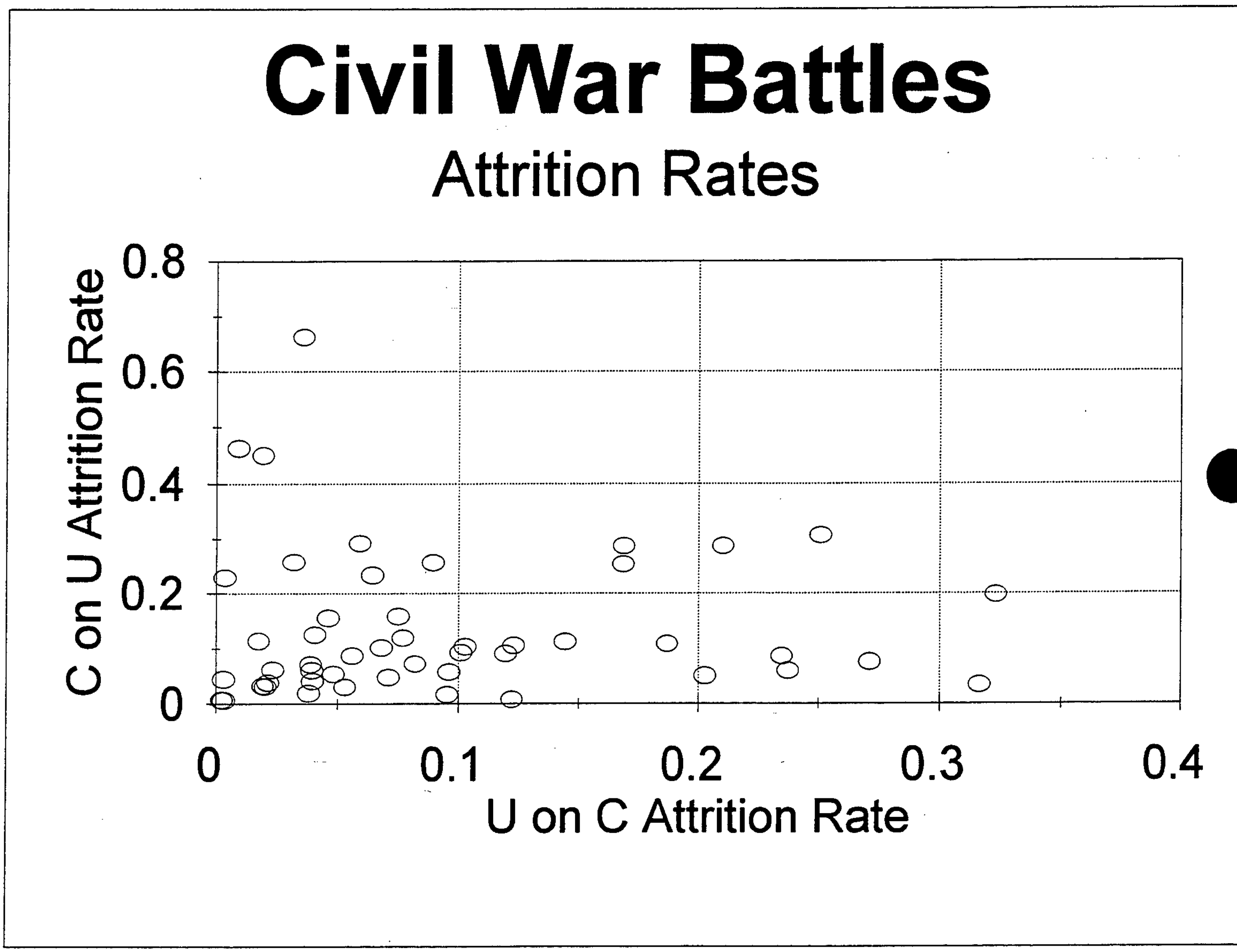


Figure XII.C.11

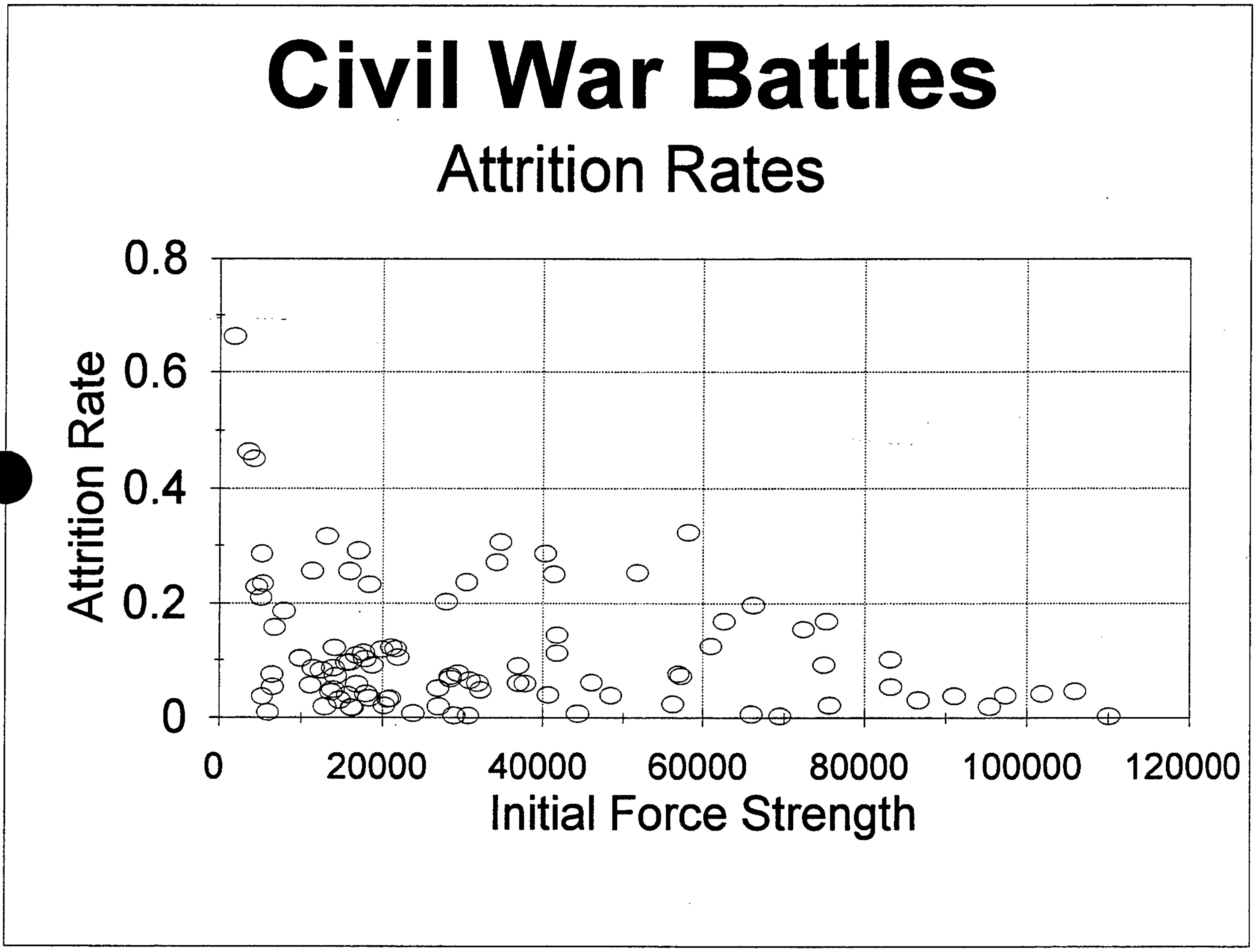


Figure XII.C.12

To determine if there is a pattern lurking in these data, we calculated the frequency distributions of the attrition rates. These are shown in Figures (XII.C-13) and (XII.C-14) for Union and Confederate, respectively. The Union distribution could conceivably be Gamma, but what is striking is the Confederate distribution. While it has no evident form, it is clear that the Confederate forces were considerably more likely to fight fiercer than the Union forces. This explains the average values from the linear regression - the Union fought more consistently than the Confederacy in terms of attrition rate.

XII.D. Meeting Engagements

Who won? It is always difficult to determine the winner of a battle? In Clausewitzian terms, there is the question of whether the result was a military victory or a political victory? If it was a military victory, was it a victory at the tactical, operational, and/or strategic level? These questions are not easy to answer and discussion still takes place over several of the battles in our data set.[g]

In this section, we return to the outline of Weiss' article. His next topic is meeting engagements, that is, battles and engagements that are not characterized by the preselection of terrain, or of its improvement, by either side. Weiss states that there were 22 battles in his data set that met these criteria. (Actually, there were 24, but Weiss discarded two as "indecisive" in outcome per his sources."

Table XII.D.1, Average Casualty Ratio, Union/Confederate "Meeting Engagements"

Winner	Union Attacker	Confederate Attacker
Union	1.09 (4)	1.06 (8)
Confederate	1.14 (4)	1.13 (6)

He presents tabular presentation of average casualty ratios, Table (XII.D.1), and average force rations, Table (XII.D.2), where the number of instances is given in parenthesis. The number of battles for each case are shown in these Tables parenthetically. Weiss notes that:

- arrival of additional units in the battles "washed out" initial effects,

[g] As a more recent example, consider Desert Storm. Today, 1993, there is still discussion about the nature of that campaign. Clearly, it was a military victory in the sense that the objectives were achieved. The consensus seems to be that it was a military victory at least at the tactical and operational levels. Was it a strategic victory? Was it a political victory? This is much less clear.

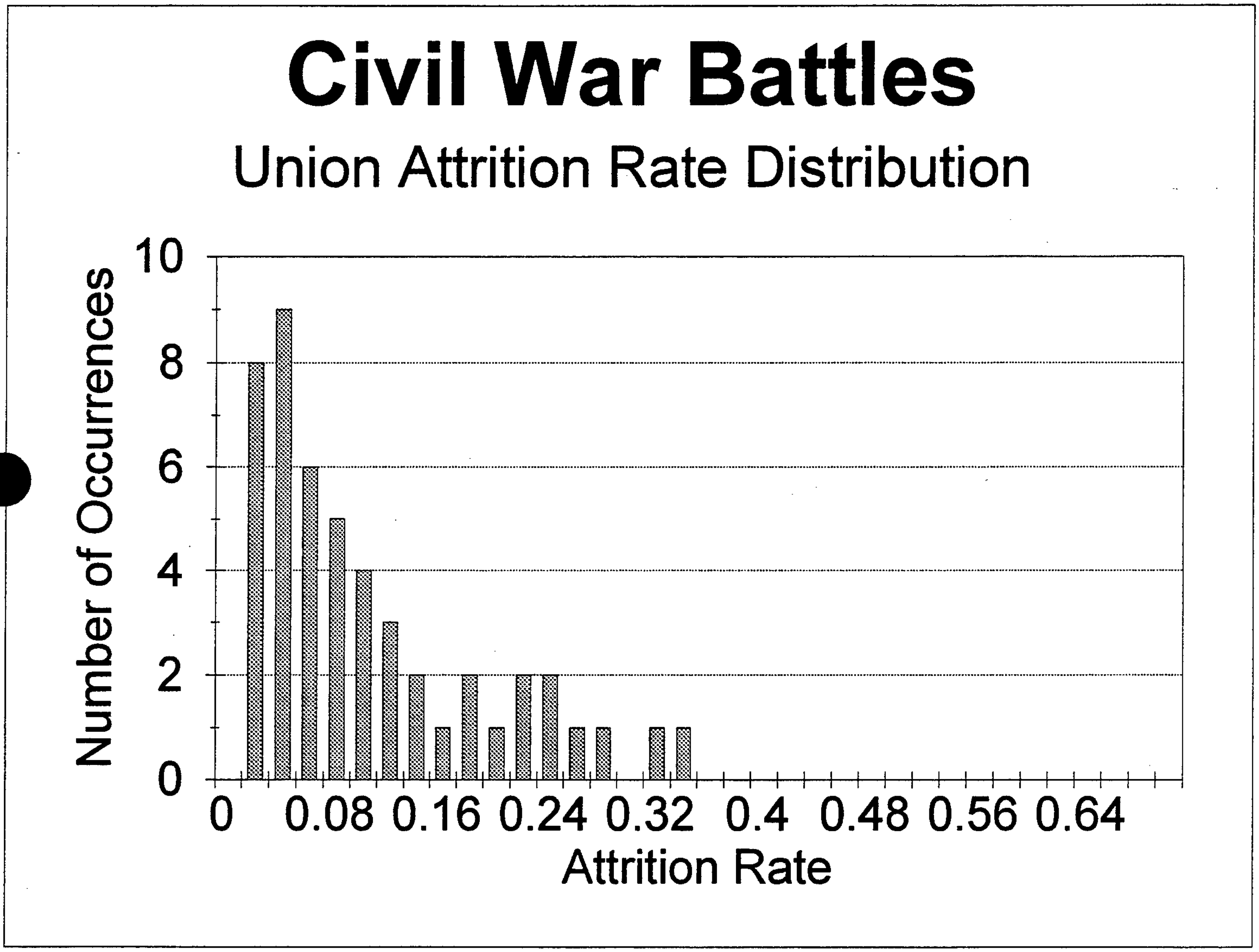


Figure XII.C.13

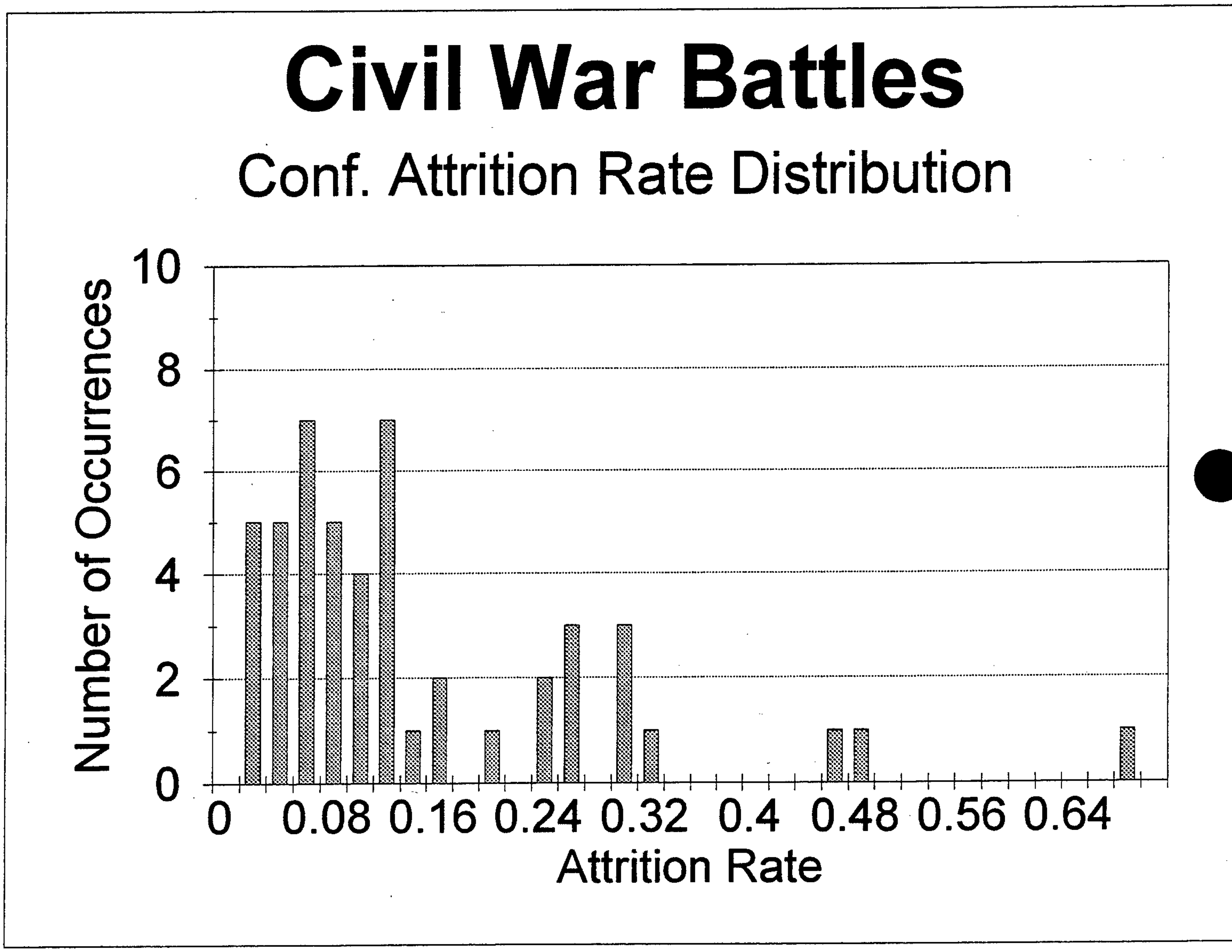


Figure XII.C.14

Table XII.D.2, Average Force Ratio; Confederate/Union "Meeting Engagements"

Winner	Union Attacker	Confederate Attacker
Union	0.62 (4)	0.70 (8)
Confederate	1.14 (4)	1.28 (6)

- each battle consisted of a series of attacks and counterattacks (This is a hallmark of meeting engagements.), and
- therefore, the designation of one said as attacker is faulty.

Table XII.D.3. Meeting Engagement Attacker Superiority Statistics

Attacker	Battles with Force Superiority	Battles attacking
Union	15	7
Confederates	8	6

Table XII.D.4. Average Force Ratio of Attacker

Attacker	Average ratio, Confederate/Union
Union	0.87 (9)
Confederate	0.96 (15)

In general, and on average, the winner had a larger force ratio, although the Confederate forces attacked with an average of 5% force inferiority. This is consistent with the thesis advanced about Confederate tactics. These data are summarized in Tables (XII.D.3) and (XII.D.4). Most critically, Weiss presents frequency data for the fraction of Union wins as a function of force ratio. These data are given in Table (XII.D.5) and shown in Figure (XII.D.1). Weiss claims that if we interpret this frequency distribution as a probability of winning, the curve is best fit by a function of the form

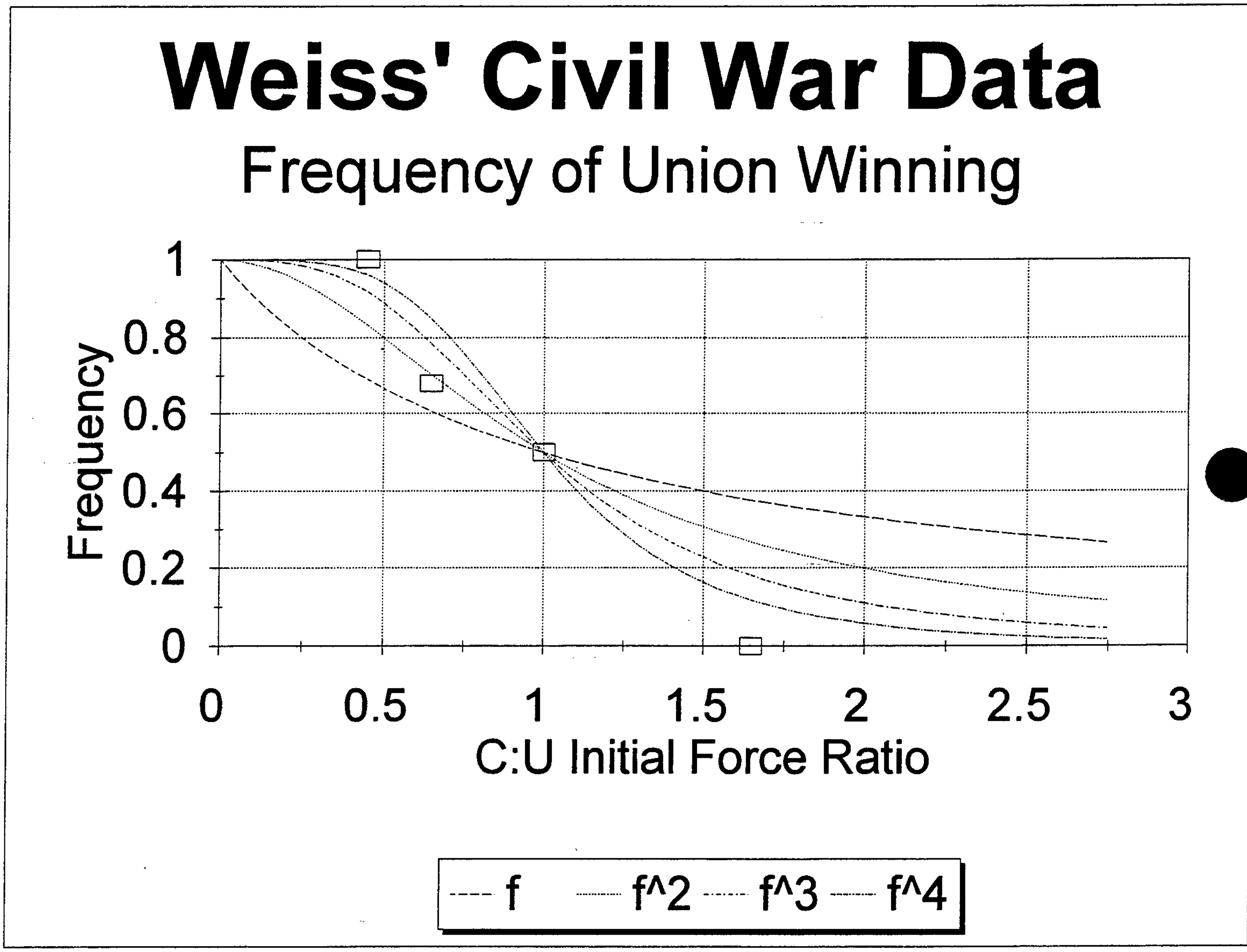


Figure XII.D.1

Table XII.D.5. Fraction of Union Wins as Function of Force Ratio

Confederate: Union force ratio	Cases	Average Fraction of Union Wins	+50% Confidence limit	-50% Confidence limit
0.40-0.49	3	1.00	1.00	0.63
0.50-0.79	11	0.68	0.80	0.54
0.80-1.25	11	0.50	0.64	0.36
1.26-2.00	1	0.00	0.75	0.00
2.01-2.50	2	0.00	0.50	0.00

$$P = \frac{1}{1 + \mu^3}, \qquad \text{(XII.D-1)}$$

where: μ = Confederate:Union force ratio (i.e. the initial force ratio).
We may examine this by comparing the data to different equations of the form

$$P = \frac{1}{1 + \mu^n}, \qquad \text{(XII.D-2)}$$

where: n = 1,2,3,4.
These curves are also shown in Figure (XII.D.1). In principle, these data can be curve fit, except that there are only two useful data points (the second and third!) This may be seen if we rewrite equation (XII.D-2) as its inverse,

$$\frac{1}{P} = 1 + \mu^n, \qquad \text{(XII.D-3)}$$

and rewrite it as

$$\frac{1}{P} - 1 = \mu^n. \qquad \text{(XII.D-4)}$$

If we now take the logarithm of this equation, we obtain

$$\ln\left(\frac{1}{P} - 1\right) = n \ln(\mu),$$

which is a linear equation amenable to curve fit. Since the equation has no intercept, however, the linear regression is trivial - it is just

$$n = \frac{\sum \ln\left(\frac{1}{P} - 1\right)}{\sum \ln(\mu)}. \qquad \text{(XII.D-6)}$$

The student may verify this fact by consulting a text on linear regression. Note that the first data point (P = 1) cannot be used because the argument of the logarithm is zero - yielding a value of minus infinity. Similarly, the last two data points cannot be used because they have P value of zero, thus logarithm arguments of infinity - yielding values of infinity. We further note that the third data point does not contribute to the numerator since it gives a logarithm argument of one - yielding a value of zero.

If we take the values of μ as the mid points of the bins, then the value of n that we obtain using equation (XII.D-6) with the data of Table (XII.D.5) is approximately 1.84. This would lead us to believe that a value of n = 2 would be a better choice than n = 3. We must note however, that Weiss does not state in his article how he arrived at his choice of n = 3 except that it was a best fit. His consideration may also have included the confidence limits on the data, and he may have used a different fitting technique or different choices of bin value for μ.[h] This example serves to demonstrate the ambiguity inherent in the data that is available. We cannot, and shall not, state that Weiss' value of three is not valid; we may only offer that there are other results possible from simple analyses such as the example above.

[h] This technique is very sensitive to the bin value selection. If we use the lower edge of the bin value for μ, then we get a value of n ≡ 0.82, while if we use the upper edge of the bin value for μ, we get a value of n ≡ 60!

Table XII.D.6, Average Casualty Ratios, Confederate/Union "Meeting Engagements"

Force Ratio	> 1.0	< 1.0
Casualty Ratio	1.17 (10)	1.14 (15)
Winner	Confederate	Union
Casualty Ratio	1.15 (12)	1.13 (14)

Table XII.D.7, Average Per Centum Casualties

	Union %	**Confederate %**	**Average %**
Winner	11.6 (14)	12.6 (12)	12
Loser	14.2 (12)	15.9 (14)	15

Weiss also examines casualties in meeting engagements. He concludes from the data, summarized in Table (XII.D.6) that for meeting engagements, "the casualty ratio is not obviously dependent on the force ratio and that casualties on both sides tend to be equal within a factor of about 2.0." Further, casualty ratios (i.e. Loss Exchange Ratios, L_{ER}s) tend to follow a log-normal frequency distribution and from this Weiss concludes that the arithmetic averages given in the tables "are consistent with a median value of casualty ratio close to unity." This suggests that the winner's per centum casualties should be less than the loser's.[i] This is summarized in Table (XII.D.7) These results may be compared with the final:initial force strength curve fits performed in the preceding section. While those fits were performed on all data, regardless of whether meeting engagement, they are consistent with these results. The distinction between the statistics for a meeting engagement and other battles must therefore be small for this data set, and presumably for the Civil War as a whole.

[i] This is Fiske's Principle of Winning, "every contest weakens the loser more than it does the winner", see Section II.D.

XII.F. Weiss' Probability of Winning Model

Weiss then develops a probabilistic model for winning. He starts by observing (in recapitulation,) that:

1. The probability of winning seems to have a strong functional relationship with initial force ratio.

2. Casualty ratios seem independent of attacker/defender, winner/loser, and initial force ratio. The range of casualty ratio is 0.46 to 2.33.

3. The average loser casualties were 15%; the average winner casualties were 12%.

In comparison to our data set, the range of casualty ratio is somewhat larger. We cannot compare winner/loser since we have not made that distinction.

Weiss postulates that the battle commences, and during its progress, continuously assesses its ability to continue. The sole criterion for the assessment is the cumulative fractional loss to that point. This is a simple model, but Weiss prudently adopts it in preference to considering perceptions of the enemy's abilities. Thus he leaves intelligence estimates of initial strengths and relative losses to further work.

This is noteworthy as an example of problem definition. The data will support the model that Weiss has formulated. Without greater, and possibly fruitless and ambiguous, research, they will not support the more elaborate considerations of perceptions.

We may also view this model from a Clausewitzian vein. The value of intelligence is questionable (for this era?) Clausewitz even goes so far as to lecture on what information the commander should ignore. Given this, consideration of the fractional loses is the only meaningful quantifiable criterion.[j]

The fundamental assumption in this model may be stated to be that the casualty ratio at battle's end is an exchange ratio characteristic of the battle and sustained at a constant rate throughout the battle. Since the casualty ratio that Weiss uses is the F_{ER}, and since equation (XII.B-8) gives the Lanchester Attrition Theory F_{ER} as a constant, we may conclude that this assumption, and the development of the whole model, are not inconsistent with Lanchester Theory.

---

[j] I do not mean to imply here that Clausewitz would have condoned this model. He would probably have viewed it as a false application of rules that should be avoided by the good commander. It may however, be consistent with Clausewitz's Law of Numbers, Section XI.B, even including circumstances and quality of troops.

Having established that Weiss' model is consistent with Lanchester Theory within the framework of small fractional losses that we have found in this Civil War data set (and indeed, in all of our data sets,) we now proceed to develop the model along the same lines as Weiss except that we generally substitute our notation for his where they differ. He first defines the losses at time $t < \tau$, the time the battle ends as

$$\begin{aligned} \Delta U(t) &\equiv U_0 - U(t), \\ \Delta C(t) &\equiv C_0 - C(t), \end{aligned} \qquad \text{(XII.F-1)}$$

while the losses at battle's end are

$$\begin{aligned} \Delta U_f &\equiv U_0 - U(\tau), \\ \Delta C_f &\equiv C_0 - C(\tau). \end{aligned} \qquad \text{(XII.F-2)}$$

The fractional losses at battle's end are then just

$$\begin{aligned} f_U &\equiv \frac{\Delta U_f}{U_0}, \\ f_C &\equiv \frac{\Delta C_f}{C_0}, \end{aligned} \qquad \text{(XII.F-3)}$$

while the force ratios at time t are similarly,

$$\begin{aligned} g_U &= g_U(t) \equiv \frac{\Delta U(t)}{U_0}, \\ g_C &= g_C(t) \equiv \frac{\Delta C(t)}{C_0}. \end{aligned} \qquad \text{(XII.F-4)}$$

Note that Weiss' f_U, and f_C are the same as our previously defined $f_{I,U}$, and $f_{I,C}$.

Weiss then defines four probability functions:

$h_U(g_U)\, dg_U$, $h_C(g_C)\, dg_C$ = probability that Union, Confederate side gives up in dg_U, dg_C, respectively, after having sustained fractional losses g_U, g_C, respectively, and

$\phi_U(g_U)$, $\phi_C(g_C)$ = probability that Union, Confederate side continue to fight at least until sustaining fractional losses g_U, g_C, respectively.

To develop these functions, he divides fractional losses into small increments (finite differences). The probability that the Union force does not give up in the j^{th} fractional loss increment is then approximately

$$1 - h_U(j\Delta g_U)\Delta g_U. \tag{XII.F-5}$$

Assuming that the eventual decision to break off the battle is independent of all previous decisions, the probability that the Union force continues the battle through n increments is approximately,

$$\Phi_U(g_U) \approx \prod_{j=1}^{j=n} \left(1 - h_U(j\Delta g_U)\, \Delta g_U \right), \tag{XII.F-6}$$

which we may transform to a sum by writing the argument of the product as the logarithm of each term exponentiated,

$$\begin{aligned} \Phi_U(g_U) &= \prod_{j=1}^{j=n} e^{\ln(1 - h_U(j\Delta g_U)\, \Delta g_U)}, \\ &= e^{\sum_{j=1}^{j=n} \ln(1 - h_U(j\Delta g_U)\, \Delta g_U)} . \end{aligned} \tag{XII.F-7}$$

Since the assumption has been implicitly made that the arguments of the logarithms are approximately one, we may approximate equation (XII.F-7) as

$$\Phi_U(g_U) = e^{-\sum_{j=1}^{j=n} h_U(j\Delta g_U)\, \Delta g_U} . \tag{XII.F-8}$$

This equation may be generalized from a sum to an integral, yielding,

$$\Phi_U(g_U) = e^{-\int_0^{g_U} h_U(g_U')\, dg_U'} . \tag{XII.F-9}$$

A similar equation may be written for $\phi_C(g_C)$.

Weiss next defines a constant R,

$$\begin{aligned} R &= \frac{g_U}{g_C}, \\ &= \frac{f_U}{f_C}, \end{aligned} \tag{XII.F-10}$$

as a direct consequence of assuming a constant exchange rate. Note that R is just the inverse of the F_{ER}, equation (XII.B-3). As we noted previously, the F_{ER} can be defined either way.

Because of the assumption of independence, the joint probability that neither side has broken off by time t (or losses g_U, g_C,) is simply the product of the two probabilities, thus

$$\Phi = \phi_U \, \phi_C. \qquad \text{(XII.F-11)}$$

The battle having proceeded thus far, the differential probability that the Union force then breaks off is

$$dQ_U = \Phi \; h_U(g_u) \; dg_U, \qquad \text{(XII.F-12)}$$

and the total probability that the Union force breaks off is just the integral of this equation,

$$\begin{aligned} Q_U &= \int_0^1 \phi_U \, \phi_C \, h_U(g_U) \; dg_U, \\ &= \int_0^1 \phi_C \, d\phi_U \, , \end{aligned} \qquad \text{(XII.F-13)}$$

and Weiss will subsequently show that this integral is trivially performable from the assumption of constant exchange ratio. That is,

$$\phi_C = \phi_U^a, \qquad \text{(XII.F-14)}$$

so that equation (XII.F-14) becomes

$$\begin{aligned} Q_U &= \int_0^1 \phi_U^a \, d\phi_U, \\ &= \frac{1}{1 + a}. \end{aligned} \qquad \text{(XII.F-15)}$$

If we define the probability that the Union force wins (i.e. does not ever break off,) as the complement of Q_U, then that probability is just

$$P_U = \frac{a}{1 + a}. \qquad \text{(XII.F-16)}$$

Weiss also computes the expected loss fraction in a battle, which is just the expected value of the instantaneous loss fraction,
which completes the closure of the model in terms of the initially defined terms. Note that equivalent quantities may also be developed for the Confederate side, both from

$$f_U = -\int_0^1 g_U \, d\Phi, \qquad \text{(XII.F-17)}$$

basic principles and from the symmetry imposed by the assumption. These are left as an exercise.

XII.G. Data Analysis

Weiss next addresses his data set to estimate the parameters and functional forms in his Model of Winning. To do this, he postulates the following:

- Let there be N_0 battles in the data set,
- If the Union side loses, there should be $N_0 \ \phi_U(f_U)$ entries in excess of f_U.
- There should, however, be L_C battles, for a given value of f_U, which ended at lesser value of f_U, that the Confederate side lost.
- Of these battles, $L_C \ \phi_U$ should have continued to f_U, and
- There are O_U battles in the data set that continue to f_U.

From this, Weiss estimates the probability function by equating these three types of battles,

$$O_U = N_0 \, \Phi_U - L_C \, \Phi_U, \qquad \text{(XII.G-1)}$$

from which the probability function is just

$$\Phi_U = \frac{O_U}{N_0 - L_C}. \qquad \text{(XII.G-2)}$$

An identical equation can be written for the Confederate side.

Weiss then curve fits the resulting data (using a similar technique to what we have previously described,) and obtains

$$\begin{aligned} \Phi_U &= e^{-kf_U^3}, \\ \Phi_C &= e^{-kf_C^3}, \end{aligned} \qquad \text{(XII.G-3)}$$

where k = 150, and from which it is obvious that

$$\Phi_C = \Phi_U^a, \tag{XII.G-4}$$

where:

$$a = \left(\frac{f_U}{f_C}\right)^3 = R^3. \tag{XII.G-5}$$

This allows the evaluation of equation (XII.F-20). If we first integrate by parts,

$$f_U = -\Phi g_U\Big|_{g_U=0}^{g_U=1} + \int_{g_U=0}^{g_U=1} \Phi \, dg_U. \tag{XII.G-6}$$

and note that the first term on the left hand side is zero at both limits (actually this is an approximation, $\Phi(g_U=1)$ is very small, approximately 10^{-65}, so we may safely approximate it as zero!) If we now use equations (XII.G-3) - (XII.G-5) to substitute into equations (XII,G-6), we get

$$f_U = \int_0^1 e^{-kg_U^3(1+R^{-3})} \, dg_U, \tag{XII.G-7}$$

and since, from our above introduced approximation, we may extend the integration upper limit to infinity, we obtain,

$$f_U \approx \int_0^\infty e^{-kg_U^3(1+R^{-3})} \, dg_U, \tag{XII.G-8}$$

which is exactly integrable (Appendix A, Integral (A-10)), yielding,

$$f_U \approx \frac{\Gamma\left(\frac{4}{3}\right)}{k^{\frac{1}{3}}\,(1 + R^{-3})^{\frac{1}{3}}},$$
$$= \frac{R\,\Gamma\left(\frac{4}{3}\right)}{k^{\frac{1}{3}}\,(1 + R^{3})^{\frac{1}{3}}}, \qquad \text{(XII.G-9)}$$

and similarly for f_C, from the definition of R,

$$f_C \approx \frac{\Gamma\left(\frac{4}{3}\right)}{k^{\frac{1}{3}}\,(1 + R^{3})^{\frac{1}{3}}}. \qquad \text{(XII.G-10)}$$

Weiss then subdivided his battle data by ranges of R to compare with equations (XII.F-19) and (XII.G-10). We reproduce this here as Table (XII.G.1).

Weiss next wrestles with a fundamental difficulty in his development. His equation (XII.F-19) is a function of the Fractional Loss Ratio R, while his correlation analysis of combat data resulted in equation (XII.D-1) which is a function of force ratio μ. He reasons that, since casualty ratio is essentially independent of force ratio, that there is a log-normal distribution of casualty ratio r, (i.e. the Loss Exchange Ratio, L_{ER},) strongly centered on r = 1, such that the probabilities P(R), equation (XII.F-19) and P(μ), equation (XII.D-1) are related by

$$\begin{aligned} P(\mu) &= \int_0^\infty g(r)\,P(\mu r)\,dr, \\ &= \int_0^\infty \frac{g(r)}{(1 + \mu^3 r^3)}\,dr, \qquad \text{(XII.G-11)} \\ &\approx \int_0^\infty \delta(r-1)\,P(\mu r)\,dr, \end{aligned}$$

where, by definition above, R $\equiv$ μ r. He does admit of the possibility that the dispersion of the casualty data (associated with g(r)) may conceal more subtle effects than he has assumed.

Table XII.G.1, Comparison of Data with Model Results "Meeting Engagements"

Range of R	0.22-0.69	0.70-0.79	0.80-1.11	1.12-2.78
Average R	0.56	0.73	0.94	1.91
Number of Battles	7	7	7	7
Fraction of Union Wins	0.79	0.57	0.50	0.29
Probability of Win from Equation (XII.F-19)	0.85	0.72	0.55	0.13
Average Confederate Fractional Loss	0.17	0.19	0.11	0.08
Calculated Loss from Equation (XII.G-10)	0.16	0.15	0.13	0.08

XII.H. Assaults on Fortified Lines

Weiss next turns his attention to attacks on fortified lines. In the parlance of Livermore, this category includes more than just attacks on forts; it includes any battle or engagement where the defender has prepared positions for that purpose. In this case, the attacker: defender casualty ratios always exceed one and have a wide scatter. Weiss finds however, that his previous derivations still hold, although the value of the constant k is twice as large (k = 300.) He conjectures that the reticence for accepting a given casualty fraction may be explained by several reasons:

1. *The attacker will break off because he feels he has been weakened beyond the point of overcoming the defender even if he breaks through, and*
2. *The attacker is deterred by the psychological effect of attacking a fortified position; i.e. fear of excessive losses*

coupled with inadequate knowledge of the defender's strength.

Further, based on a very small sample set (three battles,) the defender also seems willing to accept only lower fractional losses. We might speculate that since defenses are prepared to support, or are supported by, a relatively small force, the defender may perceive his force to be fragile with respect to losses. Alternately, applying what we consider to be a modern doctrine, the intent of the defender may be delay. If this is the case, then he will accept fewer losses fractionally to maintain his force's fighting ability in subsequent actions after breaking off. Sadly, research to investigate the subdivision of the fortified line battles to distinguish between absolute defensive intentions, and delaying intentions has not been done.

Weiss further notes that Confederate losses were small except when Union assaults were successful. He suggests then that a model need primarily be concerned with the attacker's fractional loss and the initial force ratio. Further, if the attacker's losses are proportional to the defender's strength, then the attacker's loss ratio will be proportional to the initial force ratio. That is,if

$$\Delta U \simeq C_0, \tag{XII.H-1}$$

then

$$f_U = \frac{\Delta U}{U_0} \simeq \frac{C_0}{U_0}. \tag{XII.H-2}$$

He presents scatter plots to support his arguments that, sadly, we cannot reproduce here because we have not divided our data set as he has. Weiss does present a linear regression result of the attacker dispersion from his data analysis, giving

$$\Delta U = 1500 + 0.2\, C_0, \tag{XII.H-3}$$

without standard deviation. He concludes that the attacker's fractional losses show greater variation in small battles than in large battles (i.e. the intercept dominates when C_0 is small.) While Weiss admits that this could reflect the uncertainties in all casualty data, we may offer another argument that stems from dealing with aggregated data.

Simply put, when the defender has larger forces, he will defend a longer line to allow his forces to fight effectively. Alternately, the larger the line to defend, the more defenders assigned to the position by higher headquarters. Regardless, the loner the line to defend, the easier the job for the attacker since he needs only break the

defensive line in one or, at most, two places.[k] He need only concentrate forces for a breakthrough at one (or two) places and engage the defender sufficiently elsewhere to preclude the defender shifting troops responsively. Since he need not take as heavy losses everywhere except at the breakthrough points, the attacker would be expected to have greater control over his casualties and thereby less variability of them.

Weiss then divides his data set on the basis of whether the total force strength (Union plus Confederate,) is more or less than 40,000. He finds that the attacker must have at least a force advantage of 1.25 over the defender. Further, for larger battles ($>$ 40,000 total strength,) the attacker was almost twice as likely of success for an initial force ratio $\geq$ 1.25 than he was for smaller battles.

He then examines what he calls the "Stabilizing Effect of Large Numbers". A review of both types of battles, meeting engagements and attacks on fortified lines, shows greater variability when total losses were small than when large. Probabilistic formulations of Lanchester Theory (to be discussed in a later chapter,) exist, but they tend to give results which are certain when the number of troops involved is small ($\leq$ 100 or so.)[l] On this foundation, he posits that the proper way to consider probabilistic combat should not be based on probabilities that an element of a force is removed from combat (i.e. killed or wounded,) but in the probabilistic variation of exchange ratio (presumably fractional?). The distribution would be a function of various factors, including terrain, weapons, movement, supply, and the commanders' abilities. He argues that such factors would average out in large battles, resulting in less variability, as we have seen evidenced in the data presented here. To quote Weiss, "In small battles, it is possible for a small force to defeat one much larger; in large battles, chance works to the gambler's ruin."

XII.I. Weiss' Assault on Fortified Position Model

Weiss now constructs a model for assaults on fortified lines. He divides the assault into two phases:

Phase 1: The defender's losses are light, the attacker's given by Δx_{al}. If these are large enough, he may not overrun (breakthrough and hold,) the defended position.

Phase 2: The attacker breaks through the defender's position. Incremental loss rates for both sides

[k] Obviously, the attacker will only attack if he believes he has the forces necessary for victory; otherwise, he does not attack.

[l] This is the basis of an argument that Lanchester Theory is only applicable for small unit engagements, not for battles or campaigns. The problem arises from requiring the engagement to proceed to conclusion, assuming that to be the correct termination requirement.

are now assumed identical (i.e. the same as in meeting engagements.)

In Phase 1, the probability that the attacker will abandon the assault is[m]

$$P = e^{-k\left(\frac{\Delta x_{al}}{x_a}\right)^3}, \qquad \text{(XII.I-1)}$$

where: k = 300, from the previous section, and x_a is the attacker initial force strength.

In Phase 2, both sides take incremental losses which are assumed to be equal. Thus,

$$\Delta x_{at} = \Delta x_{al} + \Delta x_{dt}, \qquad \text{(XII.I-2)}$$

where: Δx_{at}, Δx_{dt} are the attacker's, defender's losses after breakthrough (at time t), respectively. The attacker's fractional loss at time t is then just

$$g_{at} = g_{al} + \eta g_{dt}, \qquad \text{(XII.I-3)}$$

where:

$$g_{dt} = \frac{\Delta x_{dt}}{x_d}; \qquad \eta = \frac{x_d}{x_a}, \qquad \text{(XII.I-4)}$$

and x_d is the defender's initial force strength. The probability that the defender loses can now be calculated, along the same lines as equation (XII.F-16), as

$$Q_d = -\int_0^1 e^{-k_a g_{at}^3} \frac{\partial e^{-k_d g_{dt}^3}}{\partial g_{dt}} dg_{dt}. \qquad \text{(XII.I-5)}$$

This integral can also be solved by introducing the approximation of extending the upper limit to infinity (since the k's are large!), and expanding the leading term (the attacker distribution) about

---

[m] Actually, I believe Weiss mislabeled this probability. It would seem to be the probability that the attacker will not abandon the assault. This seems consistent with the way Weiss uses the formula subsequemtly in equation (XII.I-5).

$$g_d^* = \left(\frac{2}{3k_d}\right)^{\frac{1}{3}}, \qquad \text{(XII.I-6)}$$

and integrating on a term by term basis. Assuming the two k's are equal, and have the value noted previously, then the first term in this integrated expansion has the form,

$$Q_d = e^{-300(g_{at} + 0.13\eta)^3}. \qquad \text{(XII.I-7)}$$

From this equation, Weiss concludes that the quantity,

$$\begin{aligned} D &= g_{at} + 0.13\eta, \\ &= f_a - \eta f_d + 0.13\eta, \end{aligned} \qquad \text{(XII.I-8)}$$

where: f_a, f_d are the attacker, defender fractional loss ratios, respectively; should be a discriminant of success or failure in assaults on fortified lines. He has 18 data points in his data set and examines them, finding 3 successes and one partial success out of five assaults for D < 0.14, and no successes for 13 assaults for D > 0.14. From this he concludes a probability of success of the form,

$$P = e^{-kD^3}.$$

As a conclusion, Weiss notes that while fortifications vary in strength (a point we raised earlier,) the data do not support further division. Nonetheless, this equation does provide a means for estimating probability of success in a assault given the fractional loss ratios, or the attacker's fractional loss in overrunning the position (i.e. D.)

XII.J. Weiss' Wrap Up

Weiss finishes off his paper with two final sections on conclusions and suggestions for further research. In keeping with our outline, we shall summarize them here before embarking on some alternative views and comments based on our data set in the next chapter. He draws seven general conclusions:[n]

[n] I have taken the liberty of rephrasing these slightly within the context of our textbook presentation thus far, although I show these as quotes. The meaning, I hope and believe, is preserved without doing Weiss a disservice.

"1. Total losses on both sides cumulated at a fairly uniform rate after the first half year of active hostilities.°

"2. On the average, Confederate forces secured considerably more favorable local force ratios in battles than would be expected. This advantage deteriorated as the war progressed.

"3. Casualties on both sides were remarkably equal, both on the whole and in battles other than assaults on fortified lines.

"4. In battles other than assaults on fortified lines, casualty ratios appeared to be independent of initial force ratios. The probability of winning was a direct function of initial force ratio, a 2:1 advantage giving about an 0.87 chance of winning. Casualties were equal on both sides to within a factor of 2. As a result, the winner tended to have smaller fractional losses than the loser.

"5. The larger the battle (in terms of total casualties,) the smaller the statistical variation in observed casualty ratio.

"6. In assaults on fortified lines, attacker losses were proportional to defender's strength; in meeting engagements, the casualty ratio had no dependence on initial force ratio.

"7. In attacks on fortified lines, the casualty ratio showed great variability. The probability of successful attack increased with increasing attacker:defender initial force ratio. Given a favorable force ratio, the probability of success, increased with increasing total force strength. The principal determinator of success was the attacker's fractional loss ratio."

Weiss also lists four suggestions for future research:

1. *Exchange ratio versus force composition*: Bodart lists a few meeting engagements which include the number of artillery pieces on each side. These engagements demonstrate a correlation between casualty and artillery piece ratios. Investigation could be an avenue to model improvement. (Of course, access to Bodart is a prerequisite.)
2. *High fractional loss ratios on both sides of unity initial force ratio*: When initial force ratios are approximately one, the combat continued to higher fractional loss ratios than would have been indicated by the developed methodology. This indicates a potential second order effect of consider-

° We have previously commented on the variations of the data Weiss presents.

ation of relative losses as well as the primary effect of consideration of own losses.

3. *Different k factors for assaults and meeting engagements*: A more general (possibly conjugate) theory to explain the values of k other than 150 for meeting engagements and 300 for assaults on fortified lines is desirable. Weiss suggests that the relationship between k and initial force ratio may be insightful.
4. *Large versus small battles*: Dividing meeting engagements into small and large battles indicates a possibly significant relationship between casualty and initial force ratio. "For small battles, high force ratios appeared to be associated with low casualty ratios (Lanchester's Square Law)." For the large battles, the reverse happens, high initial force ratios result in high casualty ratios - a side loses strength in proportion to strength committed, regardless of enemy force strength. Lumping all together produces a uniform distribution. (Weiss refers to his figure, ours shows a different shape than his.)

The latter phenomenon noted by Weiss has been recognized by others, who he references. He notes that while vulnerability increases with force strength, effectiveness increases less rapidly (possibly even as the square root?), and suggests that a attrition differential equations of the form

$$\frac{dM}{dt} = -\alpha M \ln(N), \qquad \text{(XII.J-1)}$$

called the logarithmic law, may be applicable.

5. *Command structure and size of battle*: Since large forces have more elaborate (and redundant?) command structures, thus, large battles could be expected to have less dispersion in the casualty rate the force can fight to than a smaller force.This can be interpreted in the cumulative probability distribution analysis.

Weiss then concludes his article by calling for more analysis of historical data to support the derivation of fundamental understanding of combat on a quantitative basis. In this we can only agree wholeheartedly.

1. Griffith, Paddy, Battle Tactics of the Civil War, Yale University Press, New Haven, 1989.

2. McWhiney, Grady, and Perry D. Jamieson, **Attack and Die: Civil War Military Tactics and the Southern Heritage,** The University of Alabama Press, University, AL 1982. The higher attrition has also been associated with the Southron cultural and social mystique.

3. Weiss, Herbert K., "Combat Models and Historical Data: The U.S. Civil War", Operations Research, 14(5), September-October 1966, pp. 759-790.

4. e.g., Griffith, *op. cit.*

5. Phister, F., **Statistical Record of the Armies of the United States**, J. Brussel, pub., The Blue and The Gray Press, New York, as cited in Weiss.

6. Livermore, T.L., **Numbers and Losses in the Civil War in America, 1861-65**, Civil War Centennial Series, Indiana University Press, Bloomington, 1957, as cited in Weiss.

7. Bodart, G., ***Militur-historisches Kreigs-Lexikon (1618-1905)***, C. W. Stern, pub., Wein and Leipzig, 1909, as cited in Weiss.

8. Fuller, Major General J. F. C., **Grant & Lee: A Study in Personality and Generalship**, Indiana University Press, Bloomington, 1957.

XIII. "On the Red Field of Battle"[a]

XIII.A Introduction

There can be little question that North and South saw the Civil War in different terms. While some on both sides saw the reason for the war to be slavery, the outcome settled that political-economic issue for the United States. (It did not however, settle the issues of economic and way-of-life slavery.) Others saw the war as a fundamental; conflict between culture (or agriculture,[b] its root,) and civilization (industrialization.) Still others saw the war as a conflict between centralization and dispersion of governance, and by that, the degree of governmental infringement on the rights of the individual. While it is fundamentally crippling to argue for the merits of any society that treasures individual rights while denying those rights to others,[c] the fact remains that part of the conflict dealt with the question of strong versus weak central government. Both issues were, if not settled, at least adjusted by both the course and the outcome of the war.[d]

As the war progressed, the views also evolved. Southrons were defending their homes and ways of life - thus, the Sacred Cause. Northerners were defending the solidarity of the nation - thus, the Glorious Union.

---

[a] John Stewart and Gil Rubin, "Hallowed Ground", Longitude Music Co., (BMI) in The Cumberland Three, **Songs of the Civil War**. The title is taken from the lyrics of the song which is Confederate in theme and origin.

[b] At the Battle of Waterloo, Arthur Wellsley, the first Duke of Wellington, Commanding Allied Forces West against the French, Napoleon Bonaparte, Emperor (?) of France, Commanding, is reported to have replied to a remark "Good beans, Wellington." by the commander of the Scots Guards, with "Sir, if there is anything about which I know absolutely nothing, it is agriculture!" This moment bears great emotional weight in the movie, "Waterloo". I regret that I have been unable to find a more legitimate citation to confirm the actual words or timing. Nonetheless, the interchange is an interesting commentary on the soldier as a product of Civilization.

[c] The South has a long history of contention and contradiction on this issue; e.g. "Those who deny Liberty to others deserve it not themselves.", Thomas Jefferson, second President of the United States, who was himself an owner of slaves.

[d] Both sides had internal problems during the course of the war. The North could relatively easily introduce conscription at the price of public disapproval; the South, bound by "States' Rights" had a more difficult time even introducing conscription. While speculation is difficult and probably pointless, it is still interesting to consider whether, had the South triumphed in or at least stale-mated the conflict, the Confederacy would have had to become more centralized to balance its northern foe, and the Union less centralized in the wake of public sentiment over defeat and conscription.

The military forces also pursued their objectives in both similar and dissimilar ways. As we have seen in the data presented in the preceding chapter, both North and South fought in alike and different ways. The most striking of these must be the similarity in acceptance of actual casualties in a battle (albeit the South accepted a greater percentage.) This is especially striking given the disparity between the technological quality and quantity of armaments and the size of forces available. The lesson we may learn from this is that however good our technology, however strong we are in numbers, we cannot ignore the qualities of our soldiers and their commanders. These need development as well if the force is to be effective.

In this chapter, we follow the material presented in the previous one. We continue our examination of the concepts advanced by Weiss[1] in the context of our data sets and in the form of alternatives. In general however, we continue our investigation for some common view of warfare.

XIII.B F_{ER} and L_{ER}

In the previous chapter, Weiss advanced a most striking thought: that the fractional exchange ratio is a quantity that varies only slightly during the course of a battle. While we examined these quantities superficially in the preceding chapter, we need to reexamine them now in terms of their behavior during the course of the battle. Obviously, given the theme of this text, the vehicle for this reexamination is Lanchester Theory.

To begin this analysis, we first write the general n^{th} (attrition) order state solution in its integral form,

$$\alpha \int_{B}^{B_0} B'^{n-1}\, dB' = \beta \int_{A}^{A_0} A'^{n-1}\, dA'. \qquad \text{(XIII.B-1)}$$

Under normal circumstances, we would just perform these integrations directly, since they are elementary, but since we want to examine the F_{ER} directly, this would entail expanding the results of the integration (as we have done in the previous chapter.) Instead, we will perform these integrations approximately using the Trapezoid Rule,[2] yielding,

$$\frac{\alpha}{2}\left(B_0^{n-1} + B(t)^{n-1} \right) \Delta B \approx \frac{\beta}{2}\left(A_0^{n-1} + A(t)^{n-1} \right) \Delta A, \qquad \text{(XIII.B-2)}$$

from which we write the L_{ER} as

$$L_{ER} = \frac{\Delta A}{\Delta B} = \frac{\alpha \left(B_0^{n-1} + B(t)^{n-1} \right)}{\beta \left(A_0^{n-1} + A(t)^{n-1} \right)}. \quad \text{(XIII.B-3)}$$

From this we see that only for an attrition order n = 1, the linear law case, is the L_{ER} (and thereby the F_{ER},) obviously a constant throughout the course of a combat. Since we know that the battles in our data base have attrition orders of approximately two, and since the force strengths during the combat, A(t) and B(t), appear in equation (XIII.B-3), the "constancy" of the L_{ER} or F_{ER} is a question requiring additional investigation.

We already know, both from the above equation, and from our developments in the preceding chapter, that both the L_{ER} and the F_{ER} will be constants of the combat for n = 1. While it is not easily practicable to investigate this explicitly for general attrition order, it is quite easy to investigate the n = 2 case that we know from our AIDE calculations is a close approximation, at least for the Civil War data set. (Since we also know that attrition order tends to represent ferocity of combat, the n = 2 case also tends to represent a case somewhere between representative and worst for all of our data sets.) We may write the losses to the Red (Amber) force using the explicit Quadratic Law solution, equation (III.C-10), as

$$\Delta A = A_0 \left(1 - \cosh(\gamma t) \right) + \delta B_0 \sinh(\gamma t). \quad \text{(XIII.B-4)}$$

We immediately note that the hyperbolic terms are products of half arguments,

$$\begin{aligned} 1 - \cosh(x) &= -2\sinh^2\left(\frac{x}{2}\right), \\ \sinh(x) &= 2\cosh\left(\frac{x}{2}\right)\sinh\left(\frac{x}{2}\right), \end{aligned} \quad \text{(XIII.B-5)}$$

so that we may write equation (XIII.B-4) as

$$\Delta A = 2\sinh\left(\frac{\gamma t}{2}\right)\left[\delta B_0 \cosh\left(\frac{\gamma t}{2}\right) - A_0 \sinh\left(\frac{\gamma t}{2}\right) \right]. \quad \text{(XIII.B-6)}$$

We may write a similar equation for the Blue losses as

$$\Delta B = 2\sinh\left(\frac{\gamma t}{2}\right)\left[\frac{A_0}{\delta} \cosh\left(\frac{\gamma t}{2}\right) - B_0 \sinh\left(\frac{\gamma t}{2}\right) \right], \quad \text{(XIII.B-7)}$$

and from these, we may write the L_{ER} as (after some minor cancellations,)

$$L_{ER} = \frac{\delta B_0 \cosh\left(\frac{\gamma t}{2}\right) - A_0 \sinh\left(\frac{\gamma t}{2}\right)}{\frac{A_0}{\delta} \cosh\left(\frac{\gamma t}{2}\right) - B_0 \sinh\left(\frac{\gamma t}{2}\right)}. \qquad \text{(XIII.B-8)}$$

We note immediately, that the rate of the L_{ER} is half what it is for attrition. That is, while A(t) and B(t) vary as γ, $L_{ER}(t)$ varies as $\gamma/2$! This is an indication that we would expect the L_{ER} to change slower than the force strengths; initially, for small t, by a factor of 2. In fact, if we expand equation (XIII.B-8) for small t, the result is

$$L_{ER} \propto \frac{\delta B_0 - \frac{A_0 \gamma t}{2}}{\frac{A_0}{\delta} - \frac{B_0 \gamma t}{2}}. \qquad \text{(XIII.B-9)}$$

This equation indicates the reduced rate of change of the L_{ER}. (The student may want to compare this equation with an AIDE type of analysis - then the reduced rate is particularly obvious! I leave this as an exercise.)

We can , of course, examine the behavior of equation (XII.B-8) numerically, but prior to that action, it is worthwhile to continue its examination analytically. Let us rewrite equation (XIII.B-8) in the form,

$$L_{ER} = \frac{\delta B_0 \cosh\left(\frac{\gamma t}{2}\right) - A_0 \sinh\left(\frac{\gamma t}{2}\right)}{\frac{A_0}{\delta} \cosh\left(\frac{\gamma t}{2}\right) \left[1 - \delta \frac{B_0}{A_0} \tanh\left(\frac{\gamma t}{2}\right) \right]}. \qquad \text{(XIII.B-10)}$$

and assuming the quantity in the square braces in the denominator is small (i.e. essentially that γt is small,) expand that term to first order, giving

$$L_{ER} \propto \frac{\delta B_0 \cosh\left(\frac{\gamma t}{2}\right) - A_0 \sinh\left(\frac{\gamma t}{2}\right)}{\frac{A_0}{\delta} \cosh\left(\frac{\gamma t}{2}\right)} \left[1 + \delta \frac{B_0}{A_0} \tanh\left(\frac{\gamma t}{2}\right) \right]. \qquad \text{(XIII.B-11)}$$

If we now do some minor algebra, and keep only terms of tanh to the first power (consistent with the expansion of the denominator,) we get

$$L_{ER} \simeq \delta^2 \frac{B_0}{A_0} + \delta \left[\delta^2 \frac{B_0^2}{A_0^2} - 1 \right] \tanh\left(\frac{\gamma t}{2} \right). \qquad \text{(XIII.B-12)}$$

(If we were to expand tanh in this equation, and compare that result to a similar denominator, leading term expansion of equation (XII.B-9), the two equations would prove to be the same.)

We recall however, that the thesis was that the F_{ER} is the term which does not vary greatly according to Weiss' hypothesis. While we leave the derivation of the exact form of the F_{ER} from equation (XIII.B-8) as an exercise for the student, we do present the approximate form based on equation (XIII.B-12),

$$F_{ER} \simeq \delta^2 \frac{B_0^2}{A_0^2} + \delta \frac{B_0}{A_0} \left[\delta^2 \frac{B_0^2}{A_0^2} - 1 \right] \tanh\left(\frac{\gamma t}{2} \right). \qquad \text{(XIII.B-13)}$$

The leading right-hand-side term is (equivalently) the same as the result we found approximately in the previous chapter as equation (XII.B-8) for an attrition order of two. The right-hand-side term in square braces is essentially the state solution and thereby represents the deviation from a draw. Therefore, we may observe that the F_{ER}, to first order in the tanh, varies at half the geometric mean attrition rate[e], times a term that is proportional to the state solution. That is,

$$F_{ER} \simeq \delta^2 \frac{B_0^2}{A_0^2} + \delta \frac{B_0}{A_0} \frac{\left[\alpha B_0^2 - \beta A_0^2 \right]}{\beta A_0^2} \tanh\left(\frac{\gamma t}{2} \right). \qquad \text{(XIII.B-14)}$$

The right-hand-side term in square braces is easily recognized as the state solution.

At time t = 0, therefore, Lanchester Attrition Theory for an attrition order of two predicts that the F_{ER} will thus have the value we previously derived in Chapter XII. The value of the F_{ER} then increases or decreases during the course of the combat (with time) according to the sign of the state solution. Of course, for an attrition order

---

[e] We introduce here the term **geometric mean attrition rate** to indicate γ since it is the square root of the product of the two sides' attrition rates α and β. This terminology avoids confusion with the **root mean attrition rate** (ξ) which is the square root of the sum of the squares of the attrition rates. That is,

$$\gamma \equiv \sqrt{\alpha \beta}\ ,$$
$$\xi \equiv \sqrt{\alpha^2 + \beta^2}\ .$$

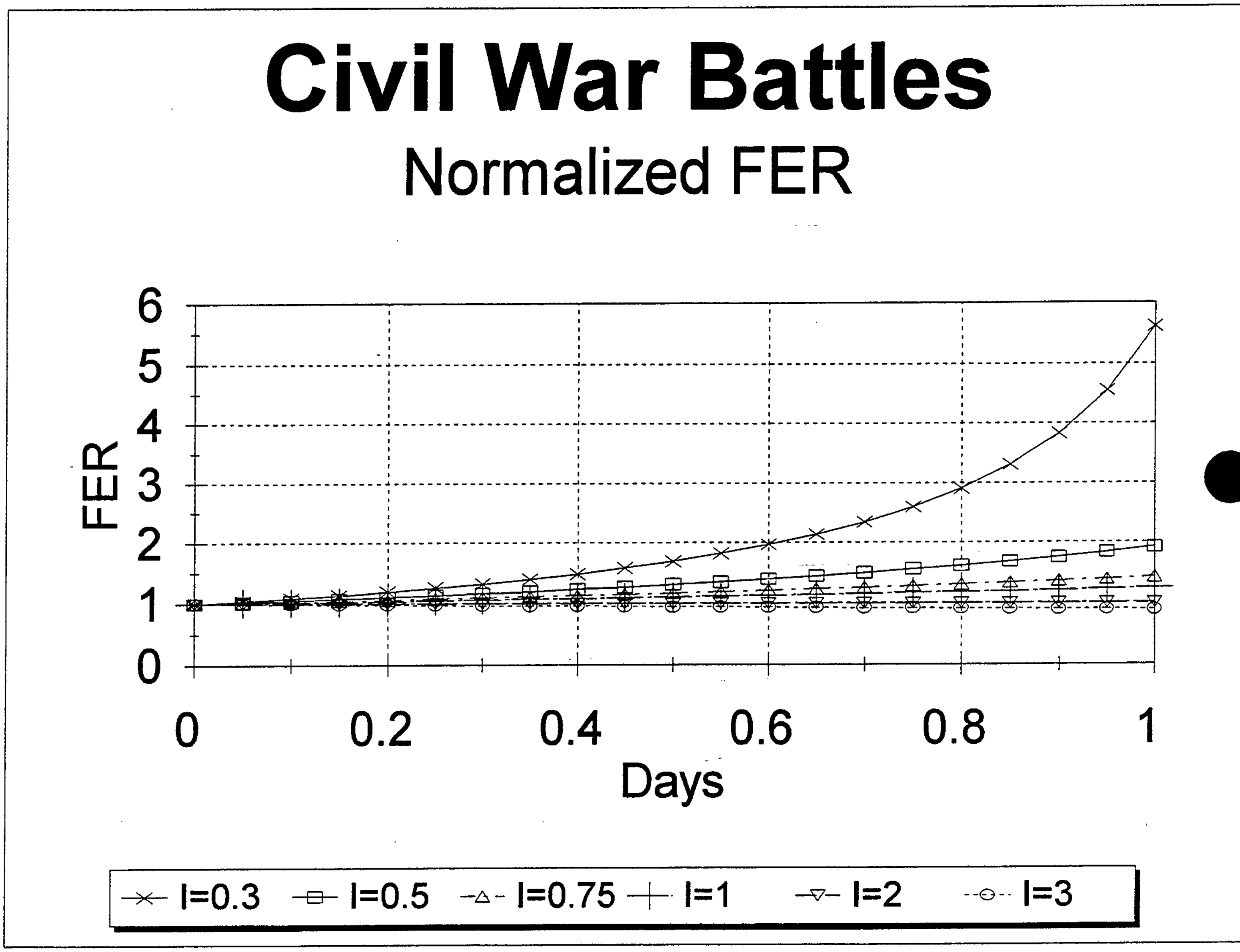


Figure XIII.B.1

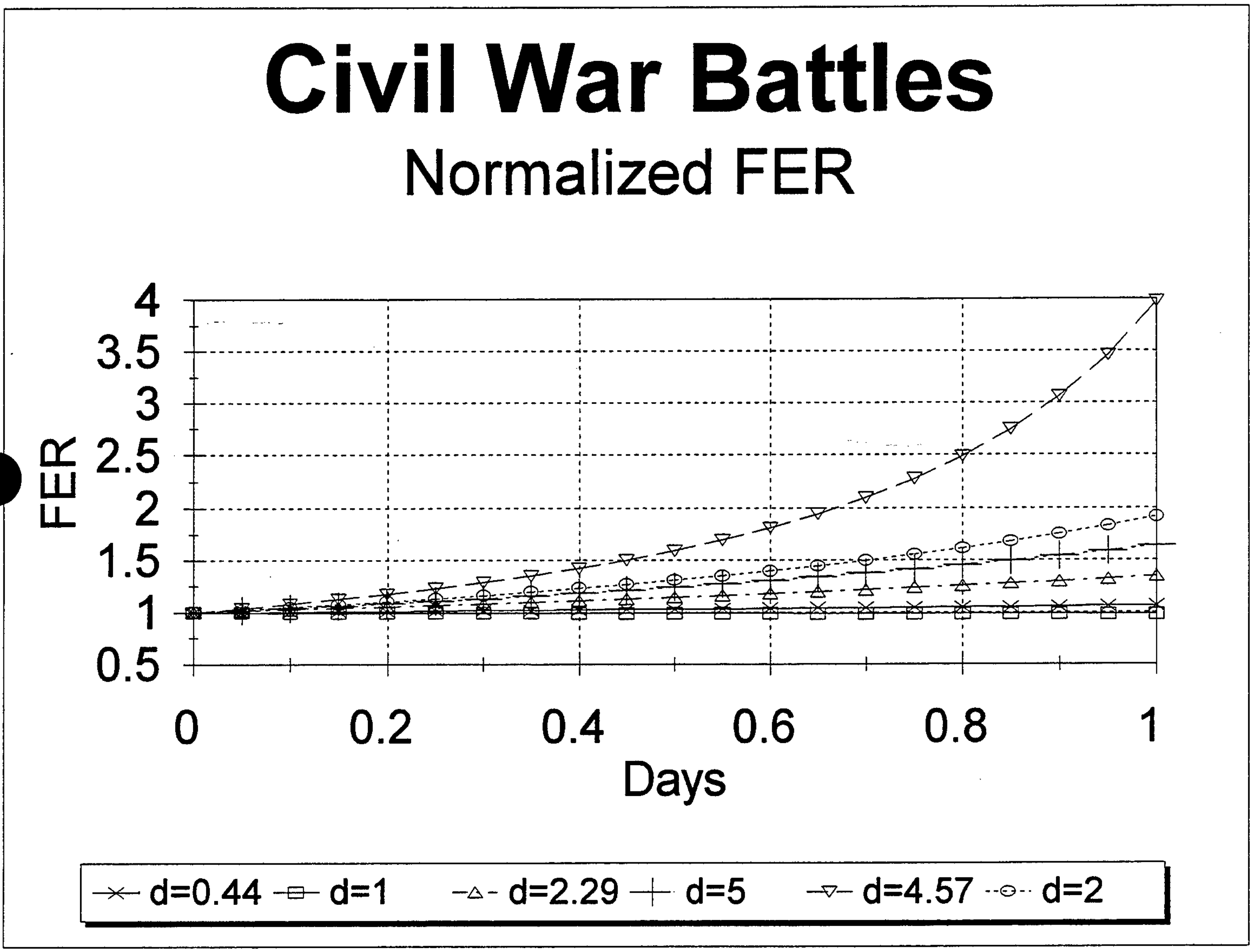


Figure XIII.B.2

of one, Lanchester Attrition Theory predicts that the F_{ER} is a constant. For attrition orders between one and two, we would expect the F_{ER} to vary, but slower than for n = 2.

How much does the F_{ER} vary? To examine this question, we perform calculations using the exact form of the F_{ER} for various initial force ratios and attrition rates. The values of the attrition rates were chosen from the Union attrition rate distribution, Figure XII.C.12. Attrition rates varied from 0.035 to 0.352. (We use the Union attrition rate distribution because it is more regular than the Confederate.) These attrition rates already include time (in days - the battle durations,) so we characterize time as fractions of a day.

In Figure XIII.B.1, we plot normalized F_{ER}'s (i.e. divided by the zero time value see equation (XIII.B-1 4),) versus time for different initial force ratios, and a δ value of 2. Over a one day combat (most of the Civil War battles in our data set lasted a day,) all of the F_{ER} curves but one vary by less than a factor of two. The curve that does vary by more than two is for an initial force ratio of 0.3. Examination of the initial force ratio statistics given in Chapter XII shows this force ratio to be an outlier.

In Figure XIII.B.2, we plot normalized F_{ER}'s versus time for initial force ratios of 0.5 and δ values of 0.44 to 0.50. Only the δ = 4.57 curve varies by more than a factor of two. This curve has the greatest value of geometric mean attrition rate and similarly to the case noted above, is an outlier.

From these calculations, we may therefore conclude that Weiss' observation that the F_{ER} does not vary by more than a factor of two during a combat is consistent with the mathematical formalism of Lanchester Attrition Theory. While it is tempting to extend this conclusion generally, we must take care in doing so. The calculations presented here, we must recall, are based on the data available on the Civil War, interpreted statistically. If we qualify the statement, then we may say that for short battles (approximately one day in duration,) with moderate geometric mean attrition rates, and initial force ratios between approximately 0.5 and 3.0, then the F_{ER} as predicted by Lanchester Attrition Theory does not vary by more than a factor of two during a combat.

XIII.C A Meeting Engagement Model

In his paper analyzing the Civil War, Weiss divides battles into two categories: meeting engagements; and attacks on fortified lines. Since his purpose is the analysis of historical data, he does not introduce an theoretical discussion of the differences between these two types of battles in terms of Lanchester Attrition Theory. Our purpose however, is the consideration of that body of theory, and we

shall therefore pause in our consideration of historical data to examine some of the mathematical applications of that theory.

In this section, we now examine a simple model of meeting engagements derived from the body of theory that we have established thus far. We turn for the basis of this theory to Section VI.H, Quadratic Lanchester Law with Reinforcements. As with all of our discussions thus far, we have limited ourselves to homogeneously aggregated combat. We shall return, in a later chapter, once we have taken up the subject of heterogeneously aggregated combat, to discuss a more elaborate model of meeting engagements.

The Quadratic Lanchester Attrition Differential Equations, as previously given in Section VI.H, are

$$\frac{dA}{dt} = -\alpha\, B + a(t), \tag{XIII.C-1}$$

and

$$\frac{dB}{dt} = -\beta\, A + b(t), \tag{XIII.C-2}$$

where a(t) and b(t) are the reinforcement rates of the Red and Blue forces respectively. The general solutions of these two equations are

$$\begin{aligned} A(t) &= A_0 \cosh(\gamma t) - B_0\, \delta\, \sinh(\gamma t) \\ &+ \int_0^t dt'\, a(t') \cosh(\gamma t - \gamma t') \\ &- \delta \int_0^t dt'\, b(t') \sinh(\gamma t - \gamma t'), \end{aligned} \tag{XIII.C-3}$$

and

$$\begin{aligned} B(t) &= B_0 \cosh(\gamma t) - \frac{A_0}{\delta} \sinh(\gamma t) \\ &+ \int_0^t dt'\, b(t') \cosh(\gamma t - \gamma t') \\ &- \frac{1}{\delta} \int_0^t dt'\, a(t') \sinh(\gamma t - \gamma t'), \end{aligned} \tag{XIII.C-4}$$

These equations are identical to equations (VI.H-1), (VI.H-2), (VI.H-13), and (VI.H-14), respectively.

We build our model of a meeting engagement from these equations. In a meeting engagement, leading portions of each force meet and engage in combat. More portions of each force arrive and enter into the combat. While one force may effectively seize a defensive posture, often due to terrain advantages, it is also common for the battle to seesaw between attack and defense for each force.[f] This means that the attrition rates for each force will also seesaw over the progress of the battle, an elaboration that our simple model does not incorporate. Instead, we simply assume an average attrition rate for each force for the entire battle. As we shall see, some insight may be drawn from even this simplified model.

In keeping with the battle outlined above, we define the initial forces that come into contact as A_S and B_S respectively, and further define the times for each force to fully deploy into the battle as τ_A and τ_B. The total force strengths available for the battle are designated by A_T and B_T, which are all related by

$$A_T = A_S + \int_0^{\tau_A} dt'\, a(t'), \qquad \text{(XIII.C-5)}$$

and

$$B_T = B_S + \int_0^{\tau_A} dt'\, b(t'). \qquad \text{(XIII.C-6)}$$

The student should note that the quantities defined here, particularly equations (XIII.C-5) and (XIII.C-6), have nothing to do with the combat itself. They merely establish the total force strengths of the two sides. Because our model of a meeting engagement starts with only part of each side in combat, we must have some accounting of the total force and the introduction of force strength into the combat. (Of course, the battle could start with all of one side deployed. The equations to be developed also include that situation although it violates our model in principle.) The reinforcement rates a(t) and b(t) are defined to be zero for t greater than τ_A, τ_B, respectively.

Before proceeding with the model's mathematical solution, it is useful to define some shortcut notation for the reinforcement terms that will occur in the solution. Accordingly we define what amount to hyperbolic function transforms of the reinforcement rates. The time dependent terms are:
and since these terms will take on constant values once all forces of each side have been deployed into the battle, the time independent terms are:

---

[f]In a heterogeneously aggregated model, we would allow the battle to be fought as a series of engagements between subcomponents or units of the two forces. These engagements could be fought alternately as offensive and defensive for each side. We shall deal with this more complex model in a later chapter.

$$\langle a(t)\rangle_s \equiv \int_0^t dt'\, a(t')\, \sinh(\gamma t'),\ 0 \le t \le \tau_A,$$
$$\langle a(t)\rangle_c \equiv \int_0^t dt'\, a(t')\, \cosh(\gamma t'),\ 0 \le t \le \tau_A,$$
$$\langle b(t)\rangle_s \equiv \int_0^t dt'\, b(t')\, \sinh(\gamma t'),\ 0 \le t \le \tau_A,$$
$$\langle b(t)\rangle_c \equiv \int_0^t dt'\, a(t')\, \cosh(\gamma t'),\ 0 \le t \le \tau_A, \tag{XIII.C-7}$$

$$\langle a\rangle_s \equiv \int_0^{\tau_A} dt'\, a(t')\, \sinh(\gamma t'),$$
$$\langle a\rangle_c \equiv \int_0^{\tau_A} dt'\, a(t')\, \cosh(\gamma t'),$$
$$\langle b\rangle_s \equiv \int_0^{\tau_B} dt'\, b(t')\, \sinh(\gamma t'),$$
$$\langle b\rangle_c \equiv \int_0^{\tau_B} dt'\, a(t')\, \cosh(\gamma t'), \tag{XIII.C-8}$$

We may now proceed to develop our meeting engagement model solution using these equations.

For times prior to the full engagement of each sides' total force strengths ($t < \tau_A, \tau_B$), the force strength time solutions are

$$\begin{aligned} A(t) &= \left(A_S + \langle a(t)\rangle_c\right) \cosh(\gamma t) + \langle a(t)\rangle_c \sinh(\gamma t) \\ &\quad - \delta\left(B_S + \langle b(t)\rangle_c\right) \sinh(\gamma t) - \delta \langle b(t)\rangle_c \cosh(\gamma t), \end{aligned} \tag{XIII.C-9}$$

and

$$\begin{aligned} B(t) &= \left(B_S + \langle b(t)\rangle_c\right) \cosh(\gamma t) + \langle b(t)\rangle_c \sinh(\gamma t) \\ &\quad - \frac{\left(A_S + \langle a(t)\rangle_c\right)}{\delta} \sinh(\gamma t) - \frac{\langle a(t)\rangle_s}{\delta} \cosh(\gamma t), \end{aligned} \tag{XIII.C-10}$$

For times greater than (or equal to) the deployment times ($t \ge \tau_A, \tau_B$), the force strength time solutions have the forms,

$$A(t) = \left(A_S + \langle a \rangle_c \right) \cosh(\gamma\, t) + \langle a \rangle_c \sinh(\gamma\, t) - \delta \left(B_S + \langle b \rangle_c \right) \sinh(\gamma\, t) - \delta \langle b \rangle_c \cosh(\gamma\, t), \tag{XIII.C-11}$$

and

$$B(t) = \left(B_S + \langle b \rangle_c \right) \cosh(\gamma\, t) + \langle b \rangle_c \sinh(\gamma\, t) - \frac{\left(A_S + \langle a \rangle_c \right)}{\delta} \sinh(\gamma\, t) - \frac{\langle a \rangle_s}{\delta} \cosh(\gamma\, t), \tag{XIII.C-12}$$

These equations are identical except that the reinforcement terms in equations (XIII.C-11) and (XIII.C-12) have reached their constant (fully deployed or committed) values. If we interpret these variables as such, that is, as variables that reach a fixed value, then the two sets of solution equations are identical.

With a simple prescription, these four equations constitute the mathematical solution of our meeting engagement model. This prescription is simply:

- if $0 \le t < \tau_A, \tau_B$, the force strength solutions are given by equations (XIII.C-9) and (XIII.C-10);
- if $\tau_A \le t < \tau_B$, the force strength solutions are given by equations (XIII.C-9) and (XIII.C-10), with $<a(t)>_c$;
- if $\tau_B \le t < \tau_A$, the force strength solutions are given by equations (XIII.C-9) and (XIII.C-10), with $<b(t)>_c = <b>_c$; and
- if $\tau_A, \tau_B \le t$, the force strength solutions are given by equations (XIII.C-11) and (XIII.C-12).

We note that regardless of the total force strengths involved, the battle cannot end conclusively prior to either τ_A or τ_B, depending on which side is concluded. Also, in the two intermediary regions, where time has progressed where one force has fully deployed, but not the other, the solution equations are intermediary in form between those above.

Before proceeding, a couple of points of discussion need to be addressed. First, all of the terms in equations (XIII.C-11) and (XIII.C-12), except the hyperbolic functions, are constants. As a result, we may form a state solution from them. That is, a function of the form,

$$\Delta_2 = \alpha B(t)^2 - \beta A(t)^2 = constant, t \ge \tau_A, \tau_B, \tag{XIII.C-13}$$

can be formed by substituting equations (XIII.C-11) and (XIII.C-12) into the above equation, carrying out all of the requisite algebra, and the result will be time independent. In a Lanchester Attrition Theory sense, this means that once all of the forces have been committed to the combat, the progress of that combat proceeds as described by simple Quadratic Lanchester attrition equations. This is not an unexpected result, but its occurrence is heartening none the less.

Of course, we could also form this state solution using equations (XIII.C-9) and (XIII.C-10) (or any A, B combination of the four equations,) but the result would be time dependent until the greater of τ_A and τ_B is reached. Such a formula would relate the inter-relationships of the force strengths, but only for times less than the minimum of τ_A and τ_B (or the relevant interval.) Only once all forces have been committed to the combat does this "state solution" become a constant and take on the proper behavior we expect of a state solution.

Second, as we have noted before, these equations are quite general. We are perfectly free to use them when A_S and/or B_S, or a(t) and/or b(t), or any reasonable combination, are zero. This allows us to study the progress of different types of combat of the general meeting engagement type. Also as we have noted, this model of meeting engagements is for homogeneous aggregation and assumes average attrition rates for the entire combat. Within these restrictions, the model permits free investigation.

Third, the restriction of average attrition rates permits us to solve the differential equations analytically and generally in the four time regions. We could introduce a scenario where the attrition rates change, and as long as they change discretely and are constant between changes, we could define time regions for all these values of attrition rates, and write solutions for each time region. Of course, this would result in a large number of solution equations, and their very number would probably confuse the issue of gaining insight.

At this point, it is useful to examine the behavior of these solutions in graphical form by actual calculation. To do this, we must assume some form for the reinforcement rates. The simplest form that we may assume is a constant rate of reinforcement, although the theory is sufficiently general that we could equally well use any mathematical form we wish.[g] The reinforcement rates then have the form

$$a(t) = \frac{A_T - A_S}{\tau_A}, t \le \tau_A, \qquad \text{(XIII.C-14)}$$

and

$$b(t) = \frac{B_T - B_S}{\tau_B}, t \le \tau_B, \qquad \text{(XIII.C-15)}$$

The reinforcement rate integrals take on particular forms,

[g] A case of particular interest that we do not address here is punctuated reinforcement where there are intervals where reinforcements arrive, and intervals where there are no reinforcements. We shall investigate this problem in a later chapter on chaos in Lanchester Theory.

Table XIII.C.1. Meeting Engagement Model Calculation Parameters

Fig. #	A_T	B_T	A_S	B_S	α	β	τ_A	τ_B
1	100	100	12.5	25	0.01	0.01	5	5
2	100	100	12.5	25	0.02	0.01	5	5
3	100	100	12.5	25	0.04	0.01	5	5
4	100	100	10	5	0.02	0.01	10	15
5	100	100	10	5	0.02	0.01	5	15
6	100	100	10	5	0.02	0.01	15	5
7	100	71	10	5	0.02	0.01	5	5
8	100	71	10	5	0.02	0.01	15	5
9	100	71	10	5	0.02	0.01	25	5
10	100	71	10	5	0.04	0.02	25	5

$$\begin{aligned}
\langle a(t)\rangle_s &= \frac{A_T - A_S}{\gamma\,\tau_A}\left[\cosh(\gamma\, t) - 1\right],\\
\langle a(t)\rangle_c &= \frac{A_T - A_S}{\gamma\,\tau_A}\sinh(\gamma\, t),\\
\langle b(t)\rangle_s &= \frac{B_T - B_S}{\gamma\,\tau_B}\left[\cosh(\gamma\, t) - 1\right],\\
\langle b(t)\rangle_c &= \frac{B_T - B_S}{\gamma\,\tau_B}\sinh(\gamma\, t).
\end{aligned} \tag{XIII.C-16}$$

The constant value forms of the reinforcement rate integrals can be formed by substituting τ_A, τ_B, appropriately in equations (XIII.C-16). These equations may now be used to perform calculations.[h]

---

[h] As an aside, if we take these reinforcement rate integrals, assume τ_A, τ_B, are sufficiently small to allow expansion of the hyperbolic functions to order τ_A, τ_B, in McLauren series (then $\langle b\rangle_s$, $\langle a\rangle_s = 0$,) and substitute into equations (XIII.C-11) and (XIII.C-12), the results are standard quadratic Lanchester force strengths without reinforcement. This confirms our earlier assertion about the "state solution". We leave this as an exercise for the student.

We present some sample calculations in figures XIII.C.l-XIII.C.10. The parameters of these calculations are summarized in Table XIII.C.1. Figures XIII.C.1 - XIII.C.3 present variations of attrition rates for forces that are totally balanced, and whose deployment times are the same, but with initial forces that vary by a factor of two. Obviously, attrition rate is the dominant factor in these calculations.

In figures XIII.C.4-XIII.C.6, again present balanced forces with 2:1 initial force ratio, but with different deployment times. While attrition rate is still dominant, it may be seen that the shorter the deployment time, the better for reducing losses. To examine this, figures XIII.C.7-XIII.C.9 present a Blue force with short deployment time, against a Red force with parametrically increasing deployment time. These amply demonstrate the advantage of short deployment time. These figures (and XIII.C.10) have forces whose normal state solution (i.e. if all forces had been deployed initially,) would indicate a draw.

Finally, figure XIII.C.10 presents the same calculations for figure XIII.C.9 except that the attrition rates on both sides are doubled. This case reverses the force strength situation. Except initially, Red never has more forces in action than does Blue. The combination of high attrition rate (even on both sides,) and long Red deployment time act to provide Blue with complete superiority. This is shown explicitly by the third curve in this figure which is the Red:Blue force ratio. Note that even after one time unit, this ratio is less than one!

We wish to emphasize that the meeting engagement model presented here will permit consideration of any form of reinforcement rate. A particular case of interest are where the average reinforcement rate (for one or both forces) is less than the attrition rate, but instantaneous reinforcement rates can be greater than the attrition rate. This type of situation is shown in figures XIII.C.11 - XIII.C.13. In this case, the reinforcement rates have the form,

$$\begin{aligned} a(t) &= 2\,\frac{A_T - A_S}{\tau_A}\,\sin^2(\omega_A t), t \le \tau_A, \\ b(t) &= 2\,\frac{B_T - B_S}{\tau_B}\,\sin^2(\omega_B t), t \le \tau_B. \end{aligned} \qquad \text{(XIII.C-17)}$$

The leading factor of 2 on the right-hand-sides of these equations are necessary for normalization (approximately! The derivation is left to the student as an exercise.) These reinforcement rates are initially zero, and are periodic with frequencies ω_A, ω_B, respectively. If we were to plot these reinforcement functions, we would find that they have bumps - they are zero at frequency-time products (i.e. ωt) equal to $2n\pi$, n = 0,1,2,..., and have maxima at frequency-time products equal to $(2n+1)\pi/2$. Operationally, we may view this as the situation when units are arriving at a battle directly from march along an avenue, are briefly formed, and then committed.

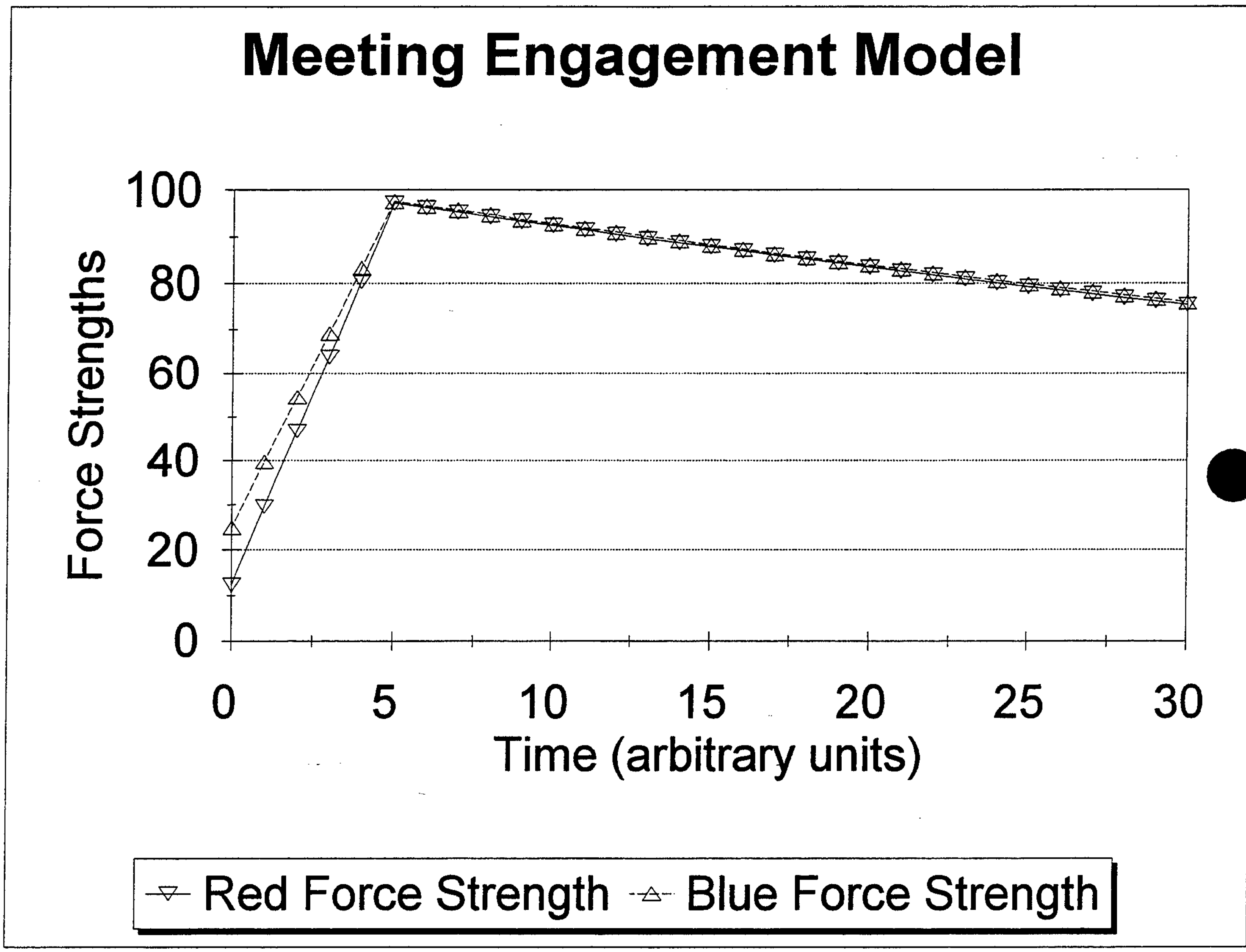


Figure XIII.C.1

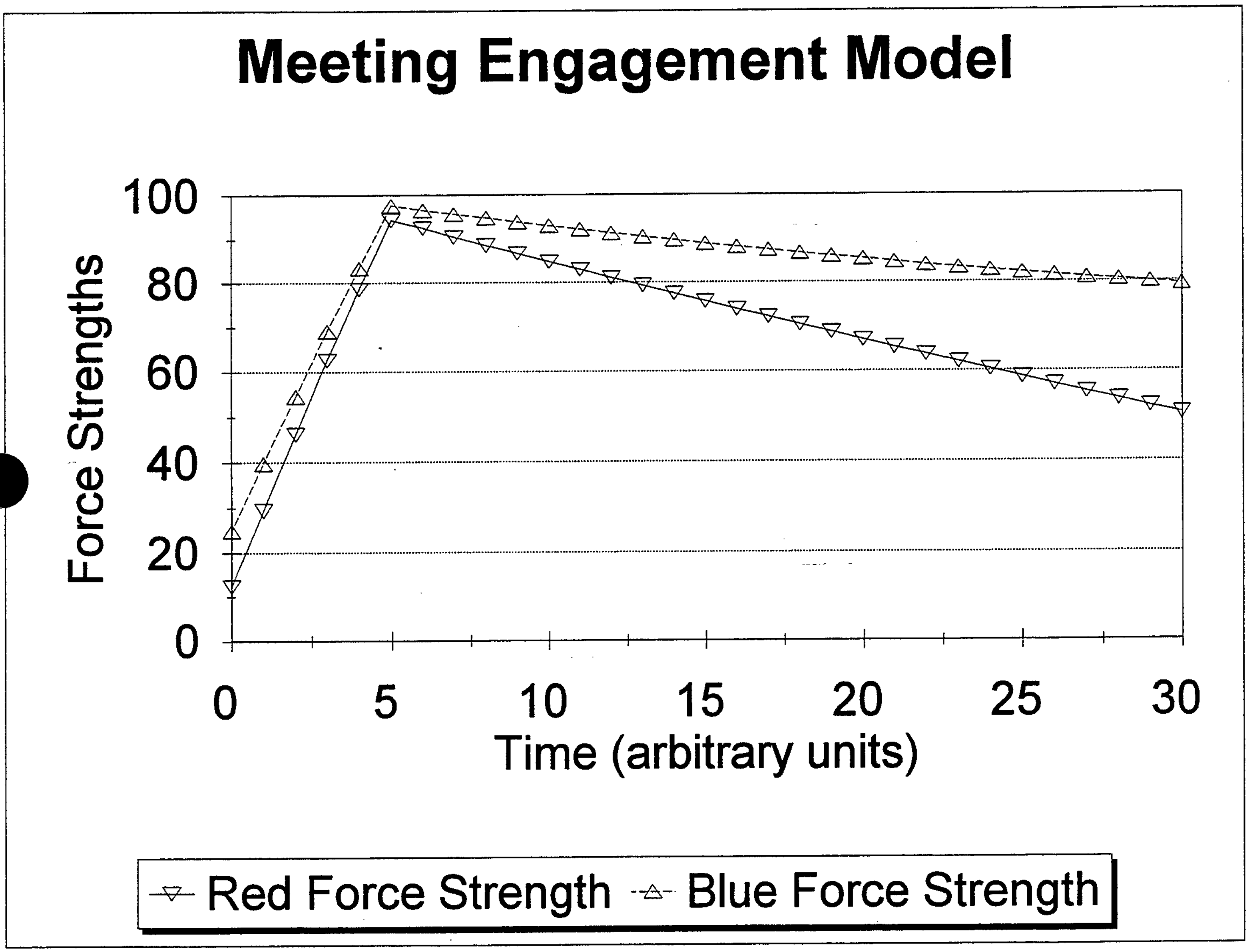


Figure XIII.C.2

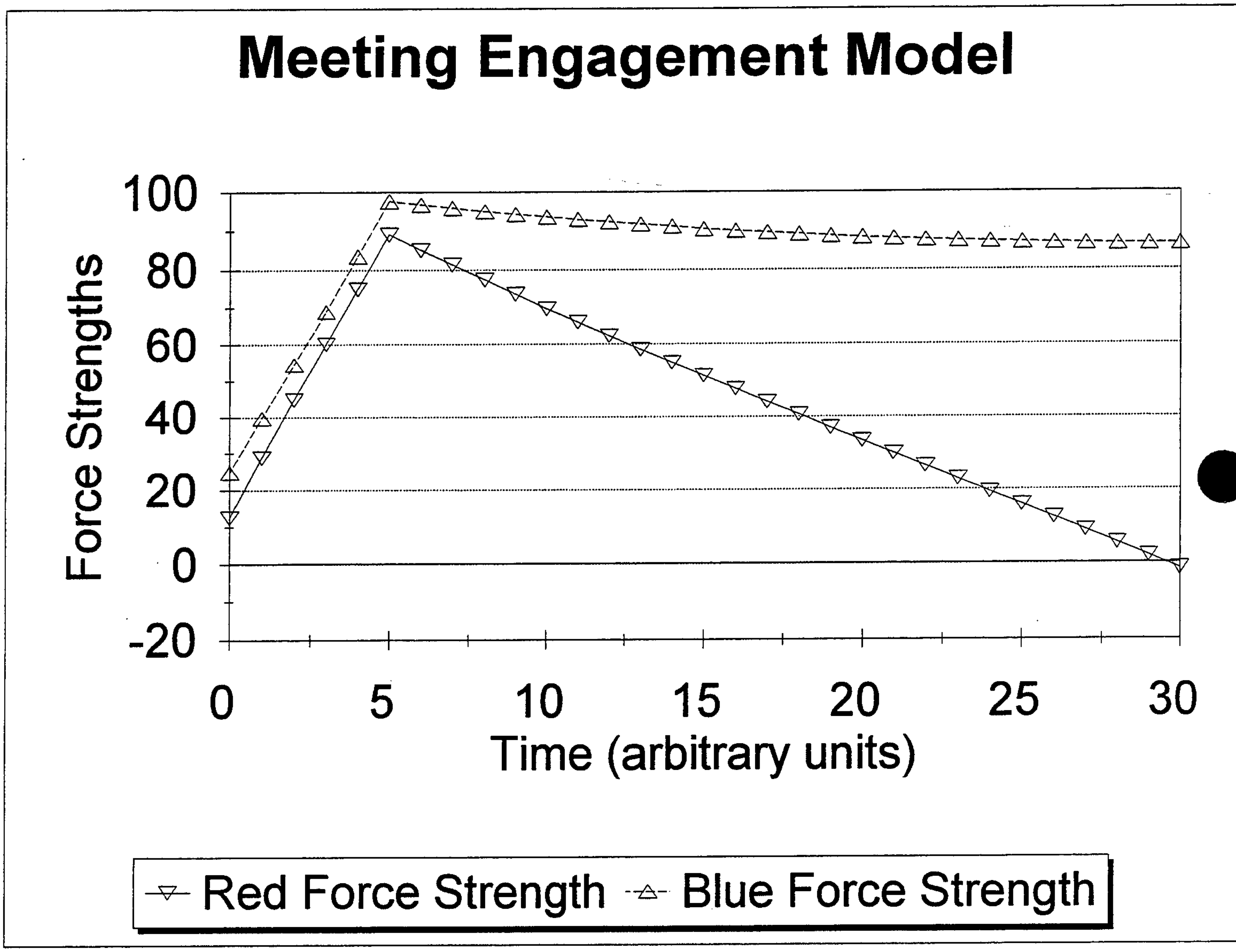


Figure XIII.C.3

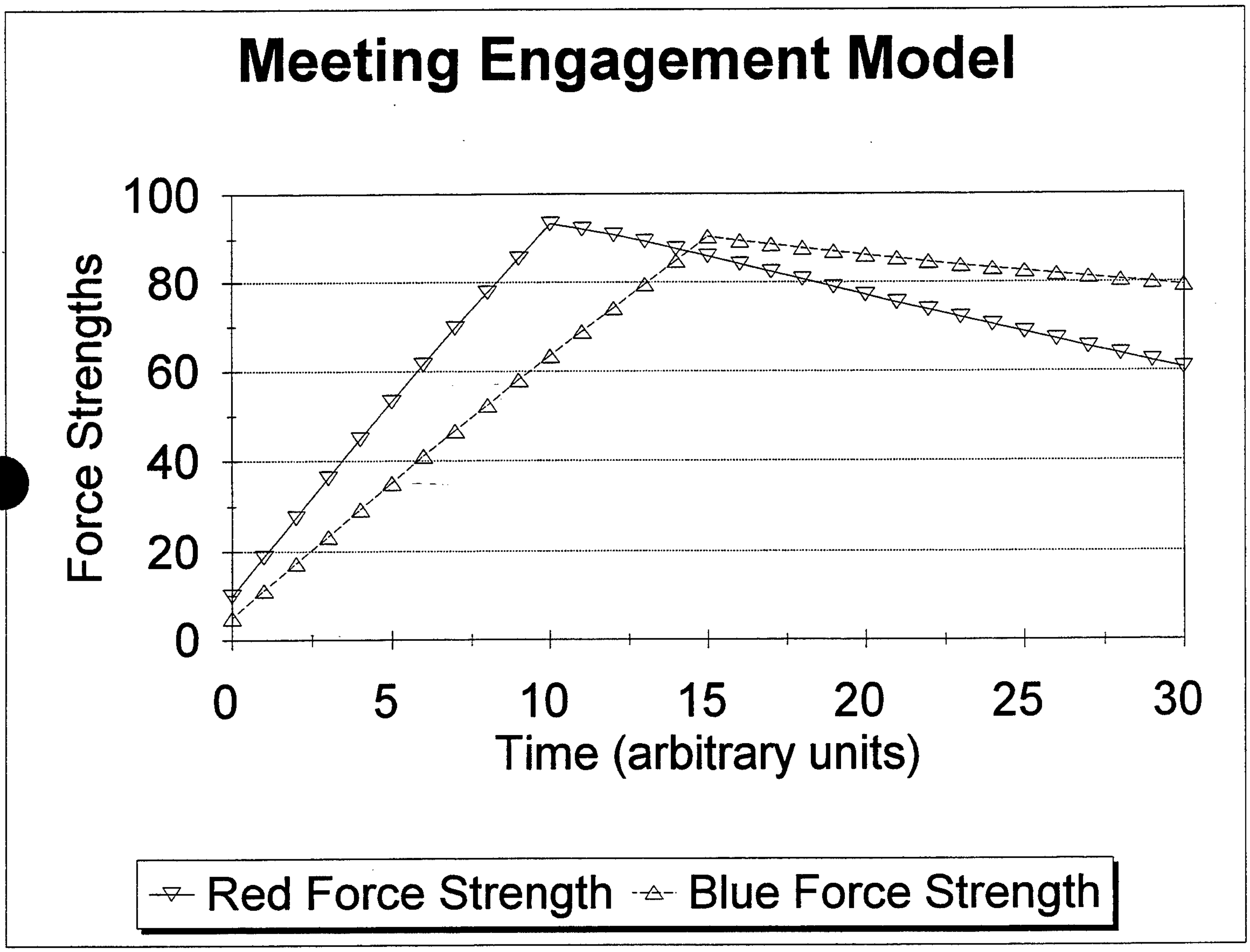


Figure XIII.C.4

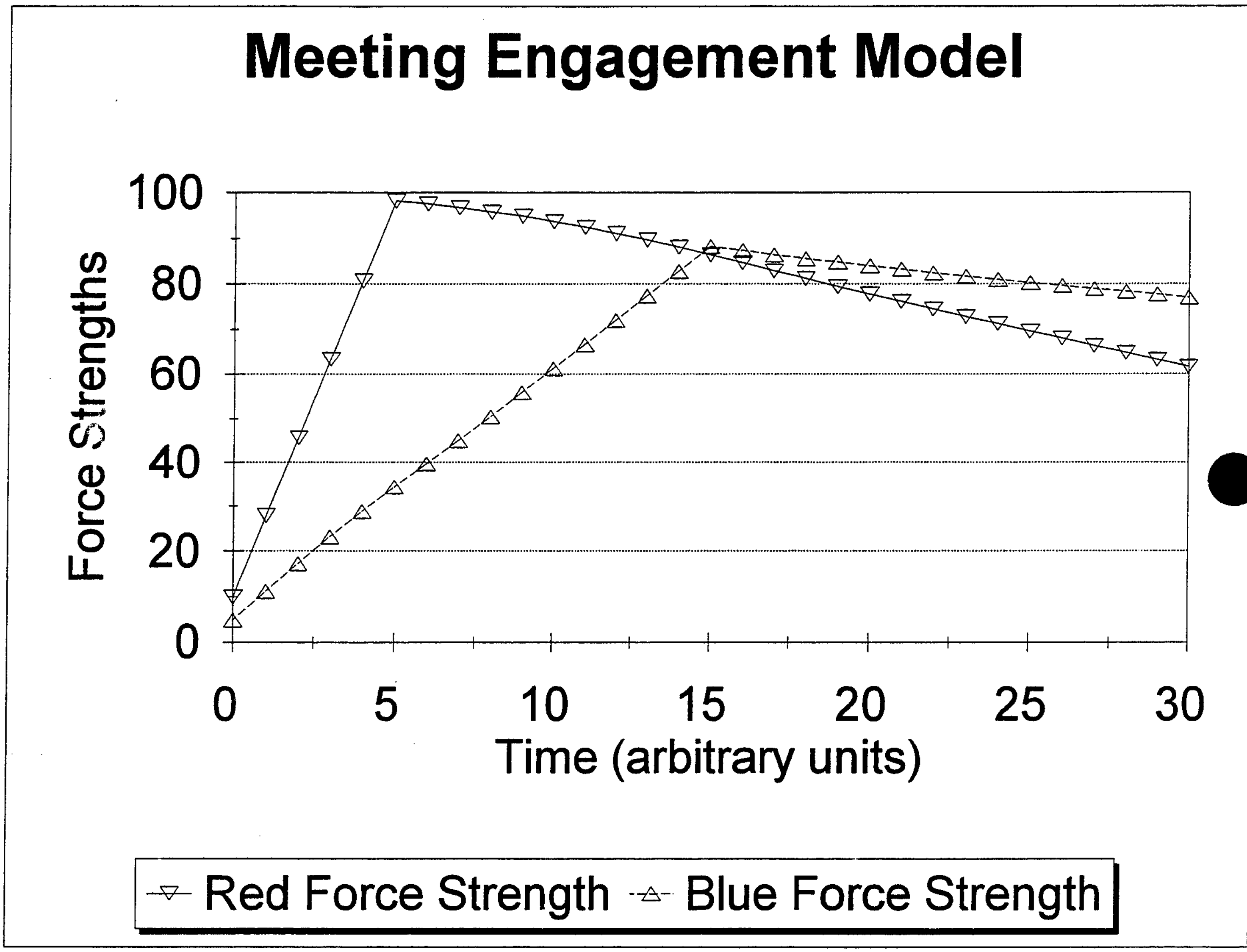


Figure XIII.C.5

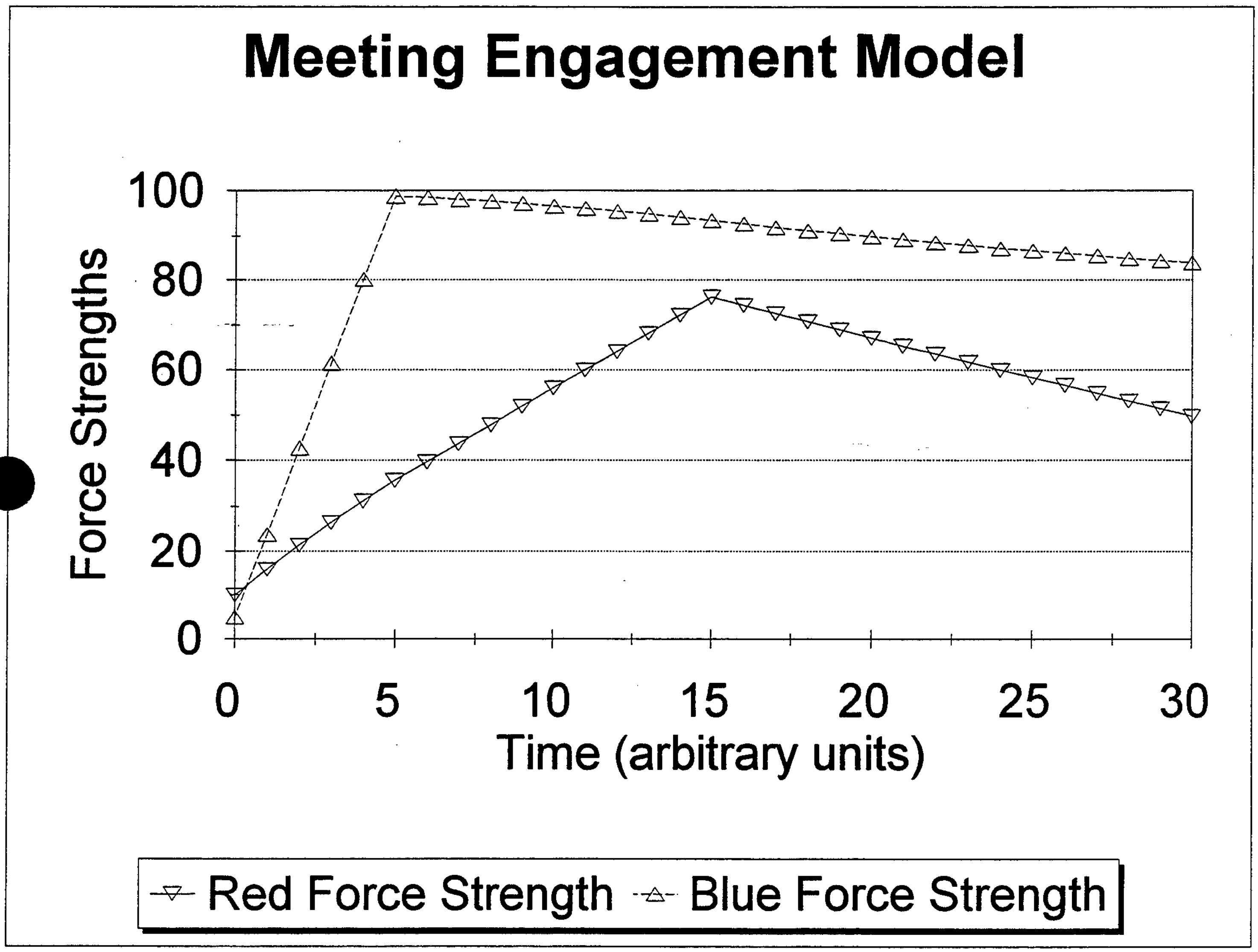


Figure XIII.C.6

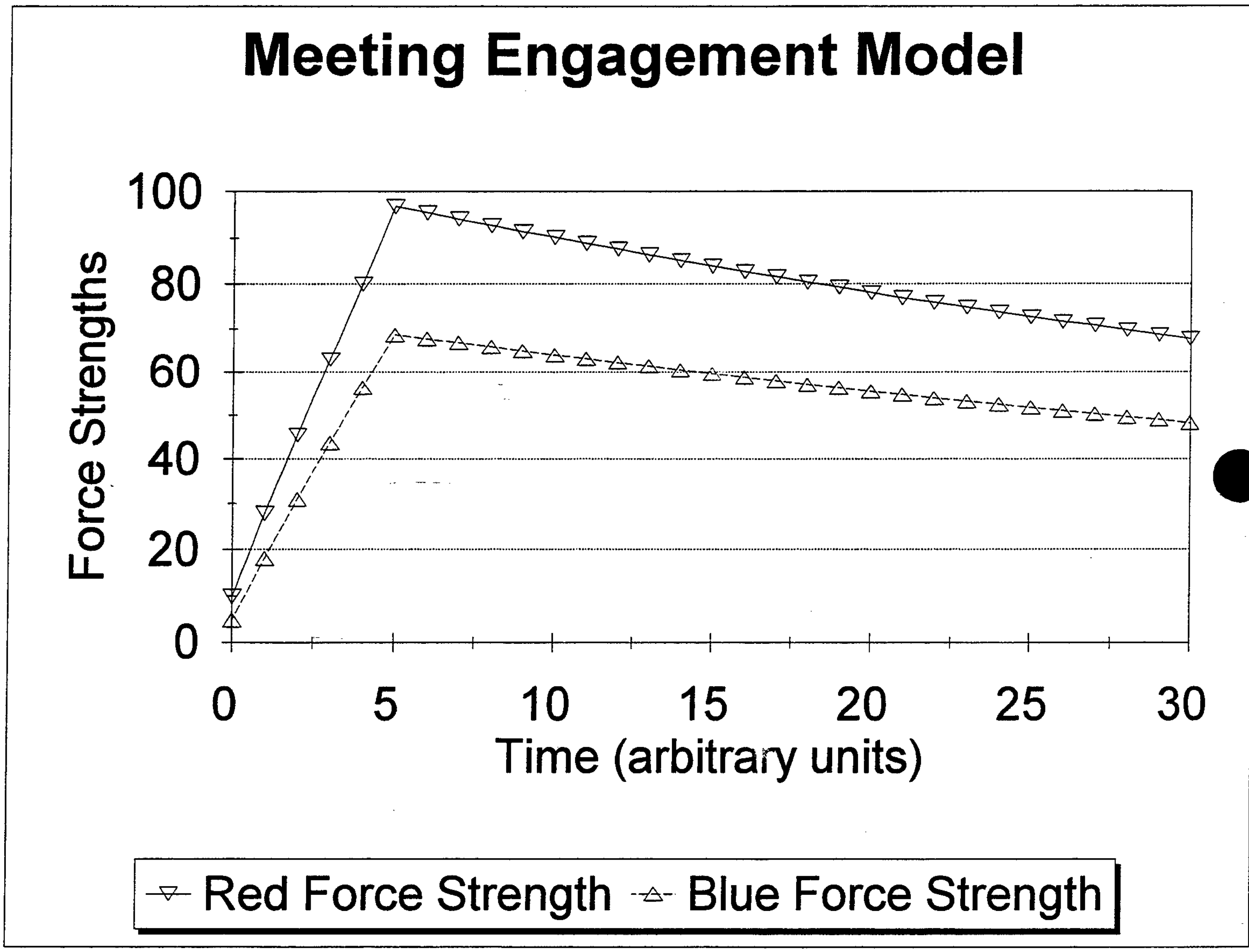


Figure XIII.C.7

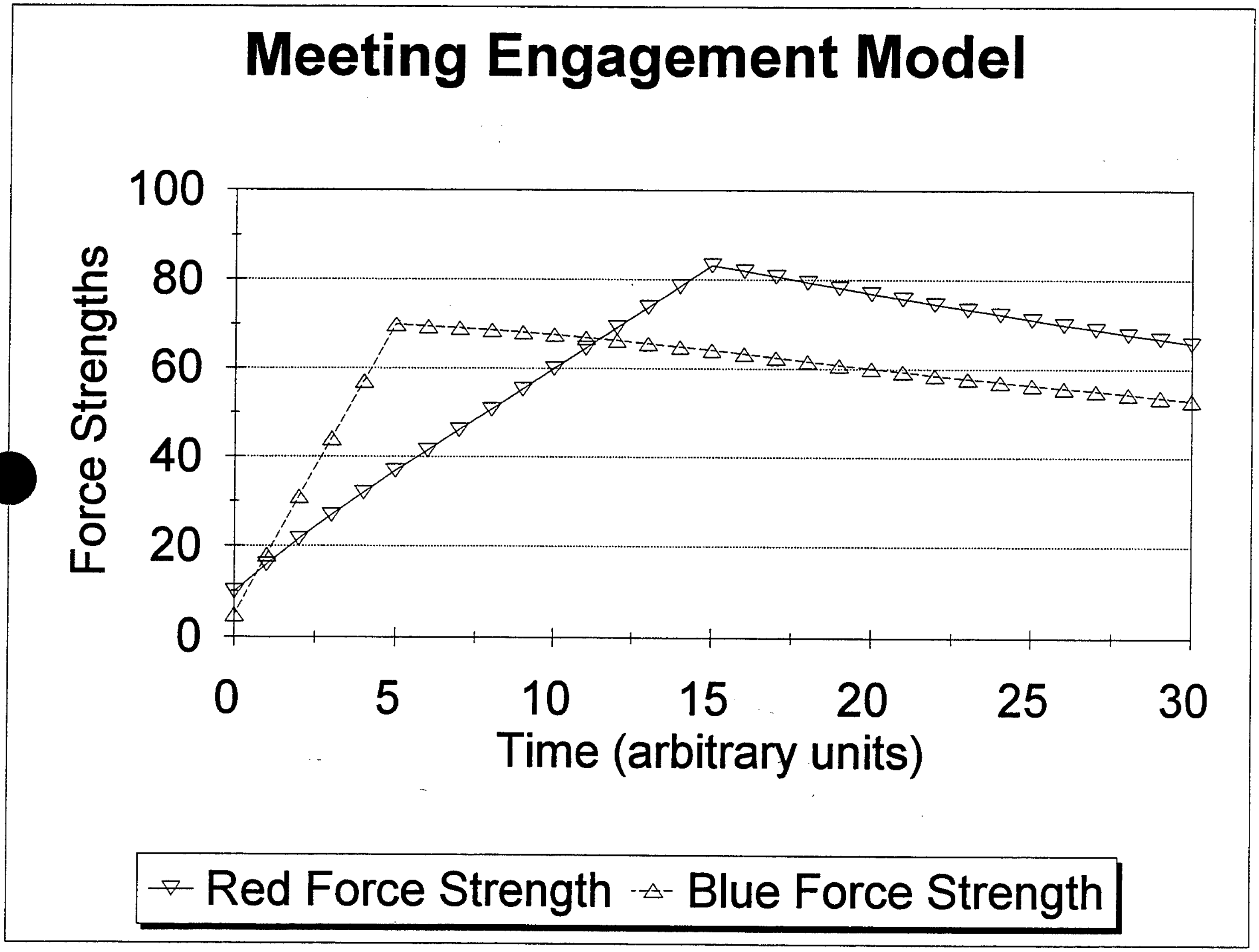


Figure XIII.C.8

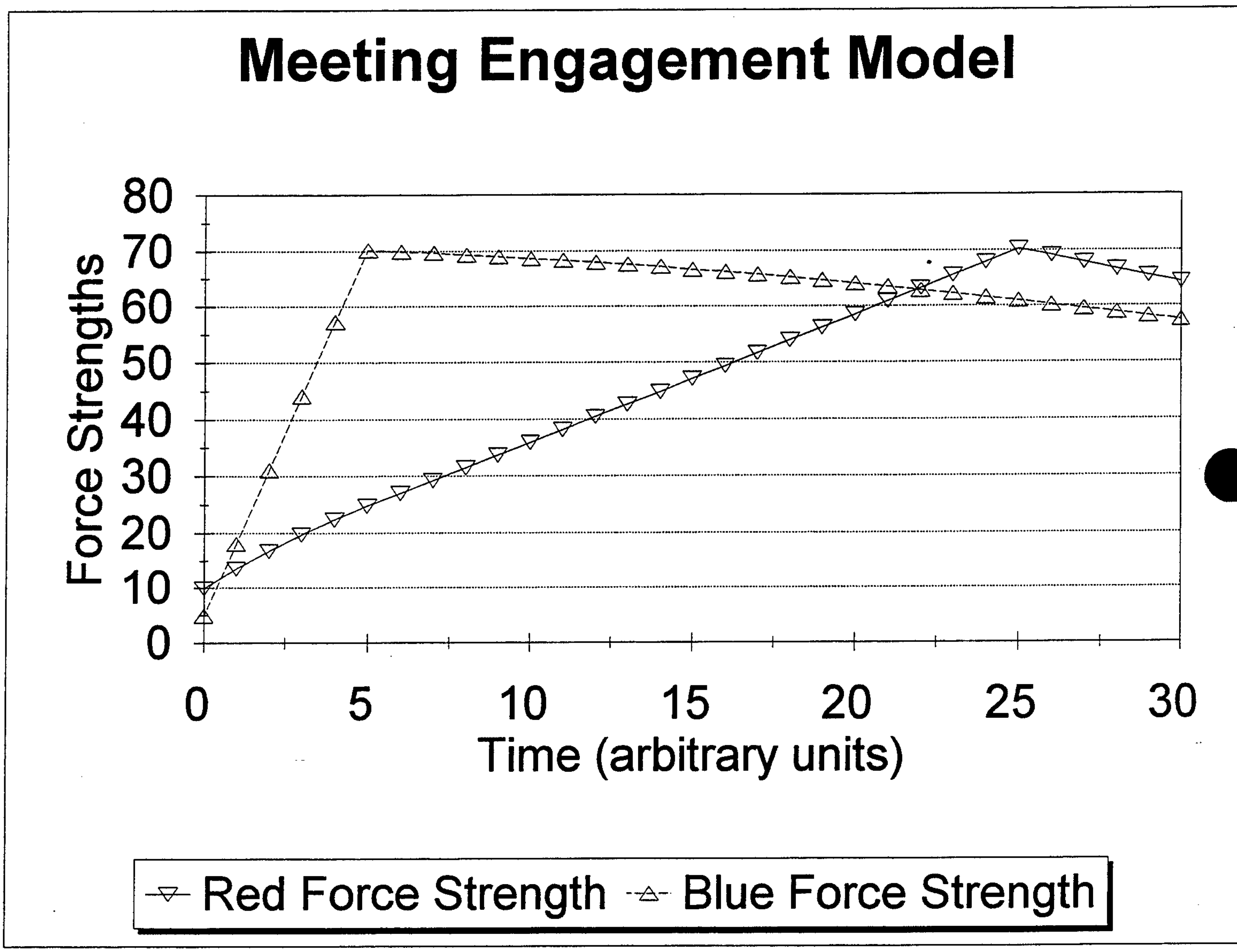


Figure XIII.C.9

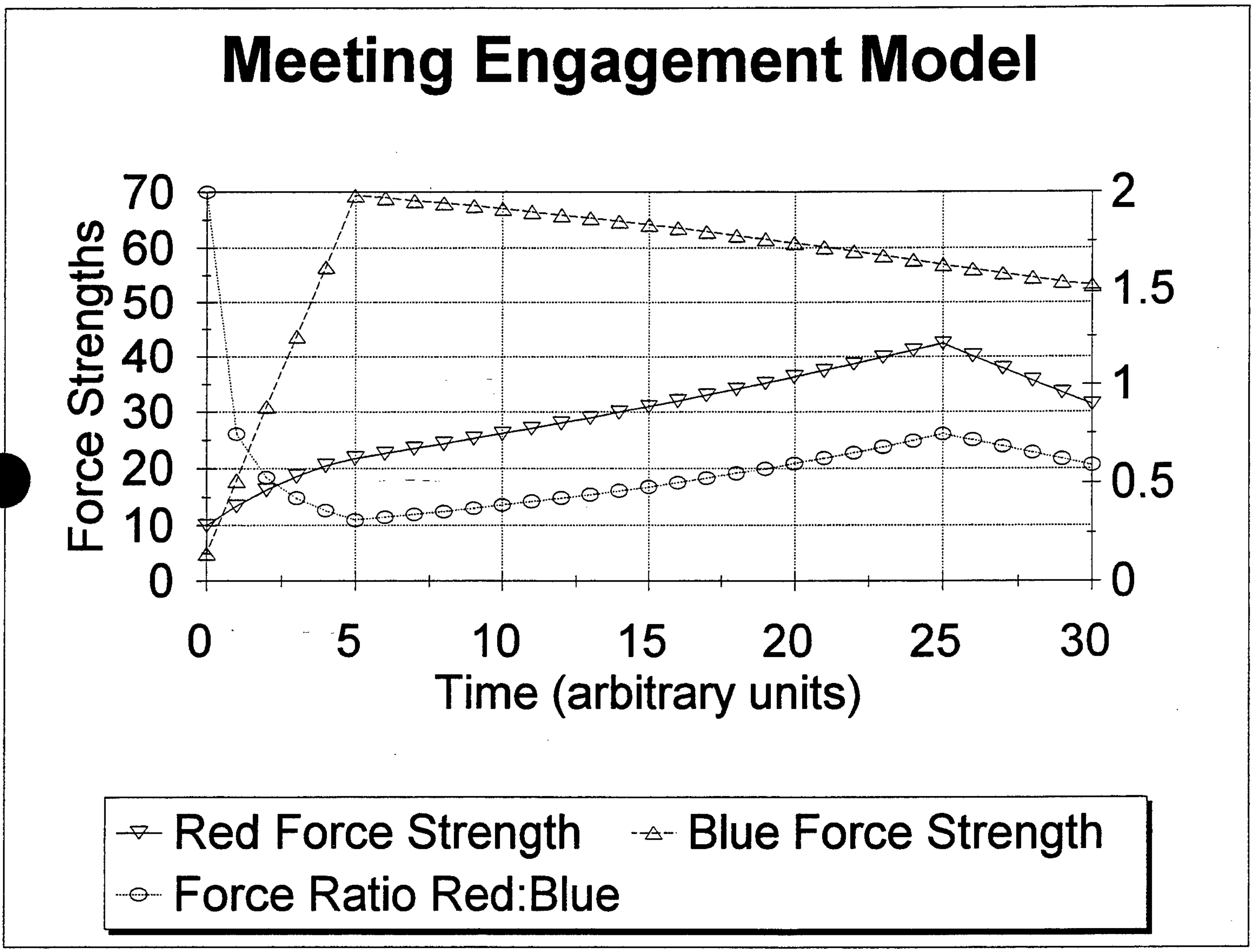


Figure XIII.C.10

Thus, the battle starts between the two starting force strengths A_S, and B_S, then a unit arrives and enters the battle, sequentially on each side. By $t = MAX(\tau_A, \tau_B)$, all forces have been committed. Obviously, there are two extremes which can occur with reinforcement rates of this form. If units arrive faster than attrition occurs, then the result is about what we saw with the constant reinforcement rates discussed above. Alternately, if units arrive slower than attrition, the effect is a series of engagements, waxing and waning as first one side and then the other gains dominance. This case is essentially the archetype of attrition which is frequently viewed as a slow grinding down of forces.

In figure XIII.C.11, we present calculations for the same parameters as shown in figure XIII.C.10, except that we use ω values of 0.3, and deployment times of τ_A = 25, and τ_B = 20. Note that for times less than 10 units, both sides suffer attrition faster than reinforcement occurs. This continues for the Red force, which never recovers, but the Blue force has seized the advantage by t = 13. The Red force is basically only throwing units into the engagement to be attrited without ever gaining an advantage. Also note the relatively small forces actually in the battle: the maximum force strengths are never above about 12% for Red, and 35% for Blue! The force ratio curve is relatively smooth here.

In figure XIII.C.12, we increase the frequencies of arrival, we $\omega_A = \pi/5$ (~ 0.63), and $\omega_B = \pi/6$ (~ 0.52). Now we see periods between unit arrival when attrition occurs faster than reinforcement, but also other periods when reinforcement is faster than attrition. This leads to the stair-step arrangement in this figure where force strength may actually decrease over short periods even prior to full deployment. Of particular note is the shape of the force ratio curve. While still smooth, it oscillates as reinforcements arrive, temporarily giving one side an advantage, until $t = \tau_B$ (the longer deployment time,) when it becomes a steady decay driven by Blue's greater attrition rate. Note also that because reinforcement is beginning to dominate attrition here, more force strength in the battle.

Finally, in figure XIII.C.13, we keep the same parameters of figure XIII.C.12 except that we reduce the Blue deployment time from 20 to 1 5. From our calculations with constant attrition rates, we would expect this to have the effect of favoring Blue, and examination of the figure shows this to be the situation. This case is intermediary between the previous two. The stair-step behavior is still present, but the reduced Blue deployment time (in this selected case,) means that Red never achieves a numeric advantage, only reaching parity in Force Strength at $t = \tau_B$. Also, the force ratio curve still oscillates with much the same form as in the previous figure, but because Blue seizes the numeric advantage because of its reduced deployment time, the oscillations after about t = 8, all have maxima less than 1 until $t = \tau_B$ when Red briefly has parity.

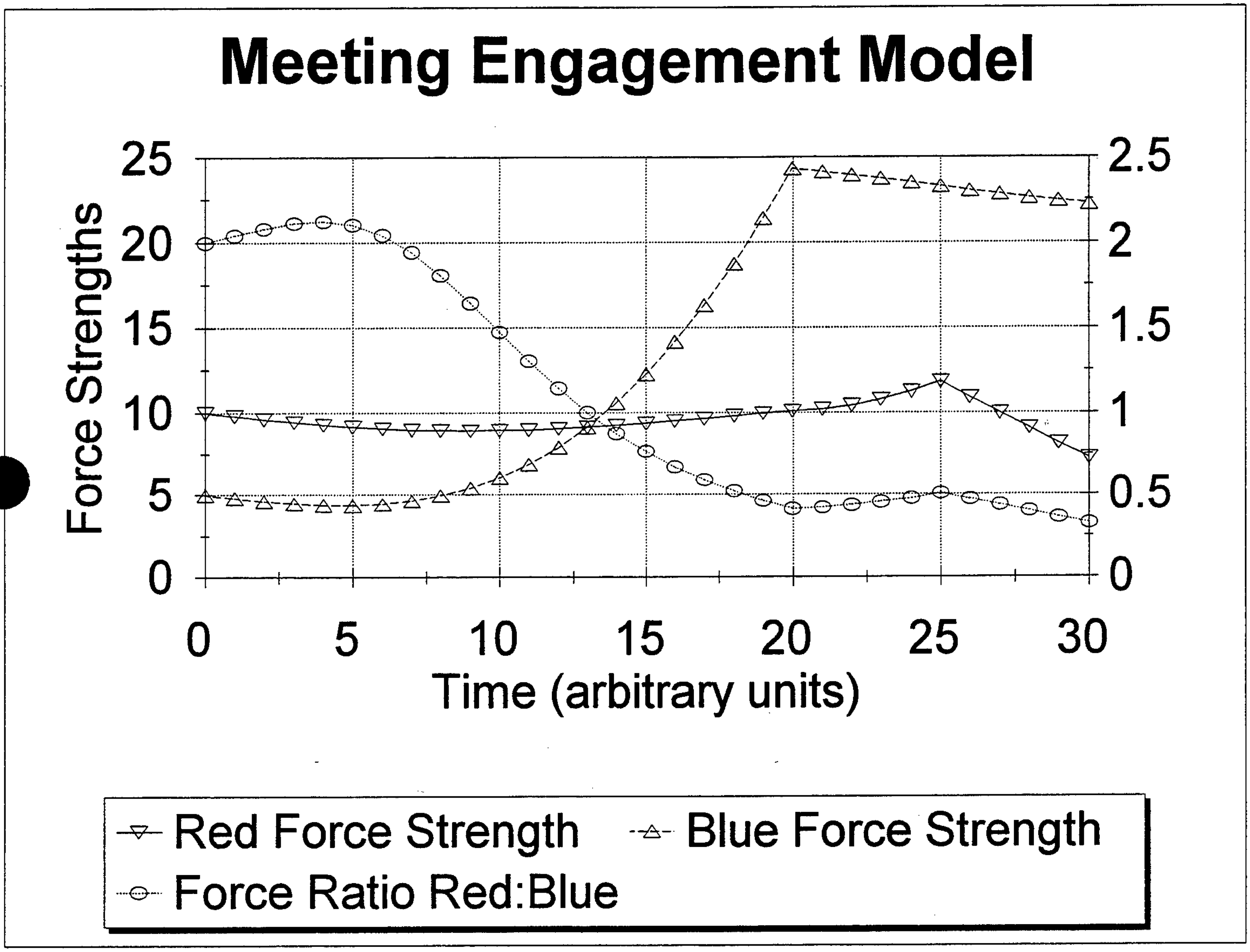


Figure XIII.C.11

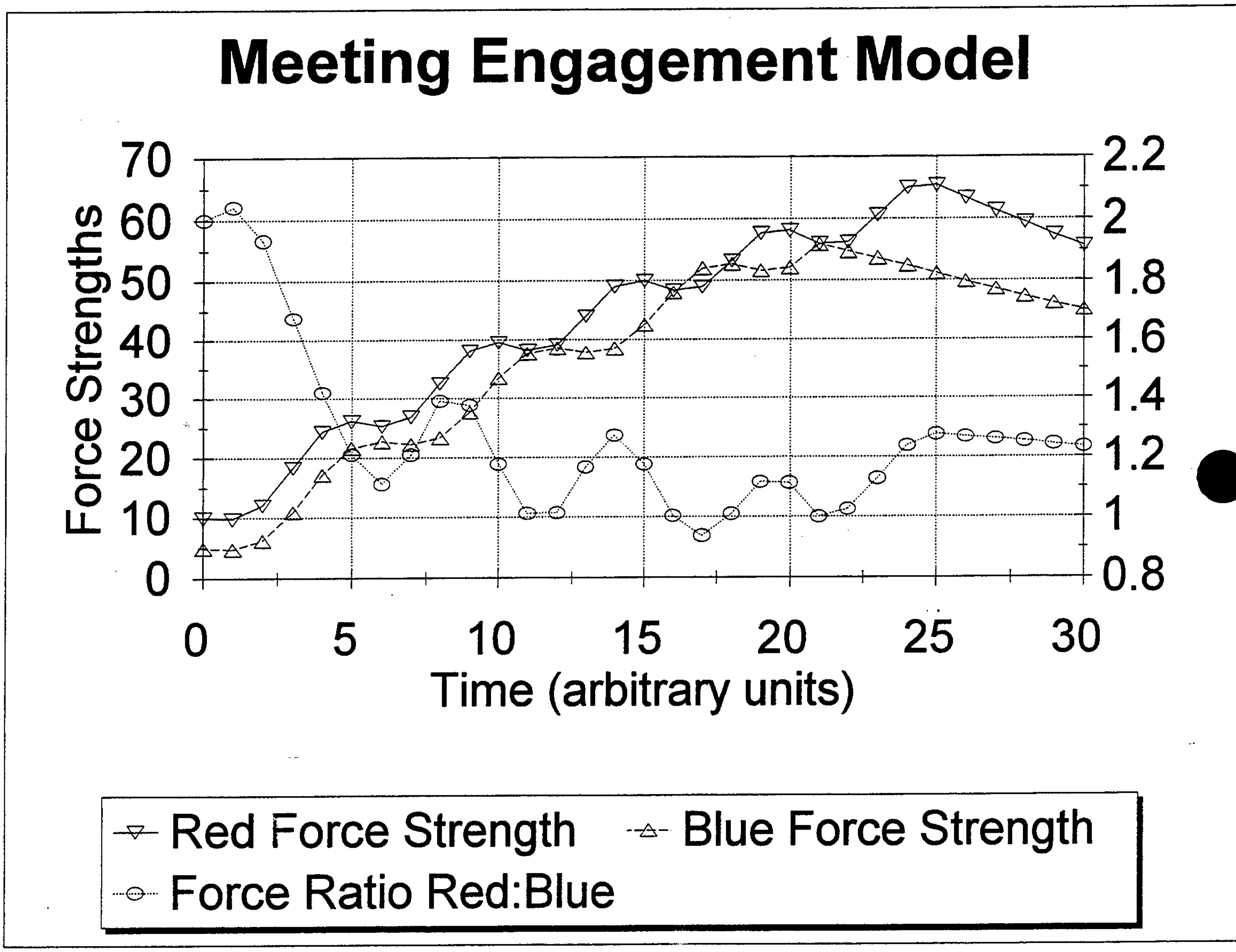


Figure XIII.C.12

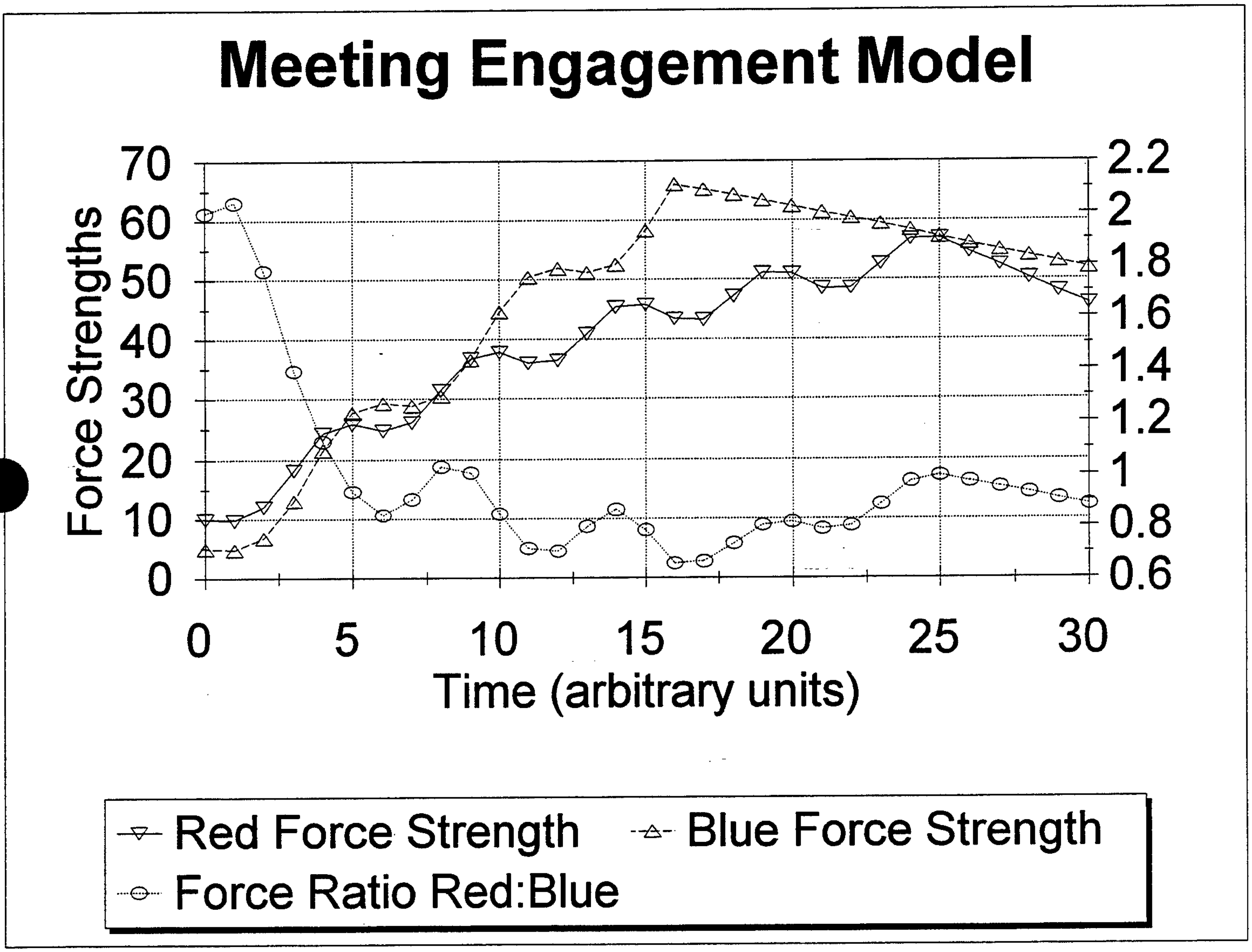


Figure XIII.C.13

This set of examples shows that the meeting engagement model here can be used to relatively generally simulate combats for a variety of conditions. The form of the model, while derived from Lanchester Attrition Theory, is far different in form than what we are used to seeing in the simple cases previously considered where all of the force strength was in the battle from its beginning. As we have seen, different insights and results may be drawn from this model than from those simple situations. In particular, one case that has been of particular interest in recent years in Lanchester Attrition Theory is that where the reinforcement rate is a function of the force ratio. We shall consider this case in the chapter dealing with chaos and Lanchester Theory.

XIII.D. Model of Attacks on Fortified Lines

In this section, we again develop a Lanchester Attrition Theory based model of a type of combat considered by Weiss. In this case, the combat type is attacks on fortified lines. To develop this model, we must suspend two of our restrictions that we have imposed thus far. We rationalize this exception to provide the student with an insightful presentation to complement the model of meeting engagements presented in the preceding section.

The first suspension is on our (already violated!) practice thus far of only considering homogeneously aggregated combat. In this case, it will be necessary to divide both the attacking and the defending forces into two separate forces, albeit that each will be homogeneously aggregated. The necessity of this will become obvious as we proceed.

The second suspension is fundamentally more important. To develop this model, it is necessary that we relax one of the assumptions central to the interpretation of Quadratic Lanchester Attrition. In particular, if we refer to Section IV.C.1, Square Law Assumptions, we need to relax assumption 2:

The units of the two forces are within weapons range of all units of the other side.

As we noted in that chapter, this assumption is really stronger than it needs to be for most considerations. It is adequate and (usually,) equivalent to only assume that the units of each force have units of the other force within weapons range that they may engage with effect. This change in the assumption is necessary to permit us to make adequate use of the Principle of Concentration to develop this model.

An attack on a fortified line is divided into two parts: the assault on and (possibly) the breakthrough of the line, and the engagement within the line. We shall treat each part in sequence. For convenience, we will assume that the Red force is the attacker, and the Blue force is the defender. The Blue force is defending a fortified line of length F_L that, if not closed, is bounded by terrain so that the Red force cannot flank it. Both Red and Blue units have a range of effective fire p with

attrition rates assumed constant inside that range, zero beyond. Both Red and Blue forces are deployed along the line of fortification. Part of the Red force will assault the line. The width of the assault force is w, and it has a speed of advance of s. We will designate the assaulting force's strength as R_1, and the remainder of the Red force strength, assumed to be deployed along the rest of the line as R_2.

Prior to the assault, Blue will have positioned his forces along the fortification, presumably with approximately uniform density. If he has time and opportunity to recognize the coming Red assault, he may have readjusted his deployment. Additionally, he may have reserved a force to reinforce the line in the area of assaults. Accordingly, we shall designate the Blue force as B_1. B_2, and B_3 to indicate the Blue force along the assaulted portion of the line, the remainder of force on the line, and the reinforcements. Except for the reinforcements, this is shown diagrammatically in Figure XIII.D.1.

Obviously, Blue would like to redistribute his forces during the assault. Equally obviously, Red does not want Blue to do this. Accordingly, Red may launch an harassing attack along all or most of the fortified line prior to the actual assault, primarily to pin down Blue's forces. Additionally, he will try to execute the assault faster than Blue can reposition his forces, or he may launch a false assault to draw forces away from the area of the actual assault. All of these embellishments are simple, given the body of theory we have already established (e.g. the harassing assault would just be modeled as a standard Lanchester engagement between all the non-reinforcing units for some period of time,) or from the parameters to be established in this model. We leave these embellishments as an exercise for the student to avoid overly complicating this presentation.

For the assault part of the engagement, we specify that the Red force has two attrition rates β_1 and β_2 on the Blue forces within the line, depending on whether the force is assaulting or pinning; that the Blue force has two attrition rates, α_1, and α_2, depending on whether the Red force engaged is fully exposed or is making use of any terrain cover or lying prone. We may assume in general that $\alpha_1 \geq \alpha_2 \geq \beta_2 \geq \beta_1$ because of the greater protection afforded in the respective engagement cases for Blue, and the greater relative effectiveness of the non-moving force for Red. (This is also why it is considered that fortified defenses take relatively smaller force strengths to be effectively defended.)

We may now proceed to write attrition differential equations for the assault. These are:

Attacks on Fortified Lines

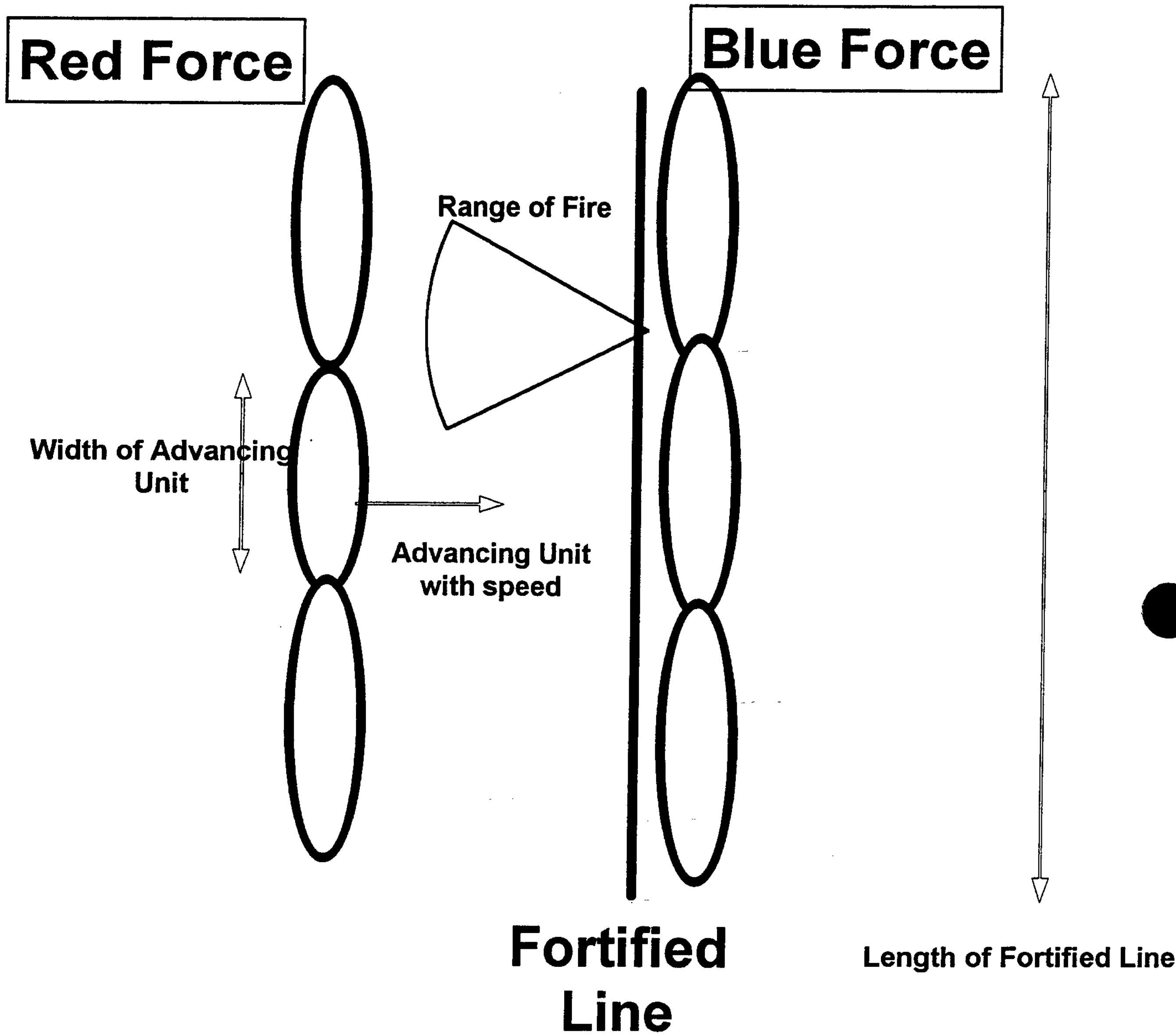


Figure XIII.D.1. Diagram of Attack on Fortified Lines

$$\frac{dA_1}{dt} = -\alpha_1 \left(B_1 + 2\, B_2\, f\, \frac{\sqrt{\rho^2 - r(t)^2}}{l - W} \right),$$
$$\frac{dA_2}{dt} = -\alpha_2\, b_2 \left(1 - 2\, f\, \frac{\sqrt{\rho^2 - r(t)^2}}{l - W} \right), \qquad \text{(XIII.D-1)}$$
$$\frac{dB_1}{d't} = -\beta_1\, A_1 + b(t),$$
$$\frac{dB_2}{dt} = -\beta_2\, A_2.$$

where the position of the assaulting force is given by

$$r(t) = r_0 - st, \qquad \text{(XIII.D-2)}$$

relative to the fortified line, f is the fraction of B_2 units that fire on R_1, as they come within range, and b(t) are the Blue reinforcements. The factor of 2 occurs because there are B_2 units on both sides of R_1, unless the assault is on the end of the line (in which case, the 2 should be removed).

In summary, these equations may be interpreted as follows:

☞ the Red assault force is engaged by the Blue force in the part of the line being assaulted plus some of the adjacent Blue force, as they can fire on the assault force;

☞ the Red "pinning" force is engaged by the remainder of the Blue force;

☞ the Blue force in the part of the line being assaulted is engaged only by the Red assault force, and received timed reinforcements; and

☞ the remainder of the Blue force is engaged by the Red "pinning" force.

Let me emphasize that this is not the only model of an assault on fortified lines that can be constructed. If the terrain permits, the Red "pinning" force, whose purpose in the assault is to prevent Blue from reinforcing the assaulted line section, could also engage that portion of the line. Further, the Red assault force only engages the portion of the Blue force in the assaulted line section, although adjacent Blue force is allowed to engage the Red assault force. We could allow the Red assault force to engage all of the Blue force that engages it, or even all of the Blue force within its range of engagement. Both of these changes complicate the form of the differential equations, and thereby the process of gaining insight from them, but these alternatives can be incorporated.

Additionally, we have not incorporated any area effect weaponry such as artillery explicitly in this model, nor have we incorporated any time at the fortified line for the Red force to breach or cross the line. This latter is a traditional problem with heavy fortifications prior to the age of gunpowder. These modifications can be done, but again, we want a simple model initially to get as much insight as we can as easily as we can, and then build up to more complicated models as we come to understand the simple ones.

Finally, we have not incorporated some very important combat process effects in this model, largely because we have not yet considered them in a detailed theoretical manner. For assaults on fortified lines, suppression and even morale are very important effects that we do not yet have a theoretical basis for consideration. We would expect that suppression is important to both sides: the Red side wishes to suppress the Blue to minimize the losses to the assault force; while Blue wishes to suppress both Red forces to maximize its lethal effect on the assault force. Similarly morale should also be important: both the Red assault force and the Blue "assaulted" force are under considerable stress; the Red force may "break" before it reaches the line, regardless of its force strength; and the Blue "assaulted" force may break once the line is breached.

Having considered all of these alternatives, at least in sketch, we may now turn to consideration of the calculational details of the model. While it is possible to solve this model analytically, despite the apparent lack of a state solution,[i] we shall solve the differential equations numerically, using a spreadsheet program.[j] The derivatives are approximated using finite differences as

$$\frac{dF(t)}{dt} \approx \frac{F(t+\Delta t)\ - F(t)}{\Delta t}. \qquad \text{(XIII.D-3)}$$

For convenience, we will take the initial position of the Red assault force to be equal to or less than the effective range of the weapons,
since they cannot be engaged at more than this range.

---

[i] We shall treat the problem of "state solutions" for non-homogeneous or heterogeneous forces in a later chapter.

[j] The solution strategy may be less than obvious on first glance. First, use the second and fourth of the differential equations of Equation (XIII.D-1) to form a second order differential equation for B_2. Solve this equation, which is an inhomogeneous, but linear equation in B_2 with time dependent coefficients using the Method of Frobenius. Next, use the first and third equations of Equation (XIII.D-2) to form a second order differential equation for B_1. Although it appears more complicated than the previous equation for B_2, it may be solved in the same manner (among other techniques.) The solutions for A_1, and A_2 may then be formed by direct integration using the third and fourth equations of Equation (XIII.D-1), and the solutions already found for B_1, and B_2.

$$r_0 \leq \rho, \qquad \text{(XIII.D-4)}$$

As we commented earlier, these differential equations do not obviously have a state solution. (Indeed, we do not expect them to have one since they have time dependent coefficients!) Unlike the model of the previous section, where there was a state solution, the timing effects were how the units arrived in the battle, and once all units were in the battle, it proceeded in a normal Lanchester manner, that is not the case here. For this model of assaults on fortified lines, timing is everything! As a result, we may expect our calculations to be dominated by those parameters that determine the timing of the engagement. (Here is one crucial point where our lack of consideration of suppression may be crucial. If suppression is fleeting unless maintained, and the forces on the assaulted line section cannot be continually suppressed during the assault, then if the assault takes too long, for whatever reason, the suppressed Blue force may recover its effectiveness.)

The first parameter we may consider is the speed of advance. For marching/striding men, a pace is approximately 30 inches, and a fast rate of pace (but not too fast to loose unit cohesion) is about 240 paces per minute (often called "double time".) Faster rates are possible,[k,3] but tend to effect unit cohesion and alignment, and if maintained over extended distances, result in lagging and fatigue. This rate of pace gives us about 120 inches per second or approximately 3 meters per second of advance. (We will try to keep to MKS units as much as possible, especially time in seconds.) Because of the relatively short range that we will associate with effective fire, we will use 4 mps for s in our calculations. Note that our results will depend on this value very strongly since it, with effective range, determine the duration of the first phase of the engagement. That is

$$t_{assault} = \frac{r_0}{s}. \qquad \text{(XIII.D-5)}$$

The second parameter that we shall consider is range of engagement. For infantry rifles of the Civil War era, ranges of 500m and beyond have been claimed, but authorities find most engagements to be fought at shorter ranges.[4] For our calculation, we shall assume an effective range of 300m, fully recognizing that this parameter, along with initial/effective range, determines $t_{assault}$ (= 75 sec.) Reducing this value will linearly reduce the value of $t_{assault}$, and thereby change the values of force strengths at the end of the assault phase of the engagement.

---

[k] Actually, this is a very fast rate of advance! The tactical systems of the Civil War era seem to proscribe "book" values of 90-180 steps per minute. Hardee's category for our fast advance would probably be a "run". We shall use this fast advance rate to compromise between Civil War and later, mechanized eras.

The last parameters that highly influence our timing are the attrition rates. We will consider the attrition rate of the assaulted force on the assault force first. In this case, the Blue force has protection and is firing on an exposed, but slowly moving (and closing) Red force. We may estimate the attrition rate by considering the probability of kill per shot and the rate of fire (as we have done previously.) Based on a basis of issue of 40-80 rounds per man, with an expected kill capacity of 4-8 men per issue (i.e. $p_k \leq 0.1$,) and a rate of fire of 4-8 rounds per minute, we may estimate a range of attrition rates. For our example, we will assume a kill every 120 sec. of engagement, which may be high, but assures that our calculations are reasonable in shape.

For the other attrition rates, we shall scale from this attrition rate based on the presented area of the target. We shall assume that the Red "pinning" force can assume positions or postures where they present approximately half the area of an upright man, so that $\alpha_2 \approx \alpha_1/2$. Further, we shall assume that the fortified line reduces the Blue man's presented area to 3/4 this, and that a moving Red firer is half as effective as a stationary Red firer. This defines the values we shall use for attrition rates in our examples.

Finally, we will assume that the fortified line is 1 km. in length, and that there are 1000 units on the Red side and 500 on the Blue. We shall vary the length of the section of assaulted line and the number of Red units in the assault, but evenly distribute Blue units over the entire line. The fraction of unit carryover to engage the assault force is assumed to be 0.5, and there is no Blue reserve.

Table XIII.D.1. Variable Assault Force Strength Results

Figure #	**Initial Red:Blue Force Ratio**	**Completion Red: Blue Force Ratio**
XIII.D-2	2	0.42
XIII.D-3	3	1.87
XIII.D-4	4	3.86

We now proceed to examine sample calculations performed using these parameters and equations. For an assault width of 100 m., we vary the Red assault force to be 100, 150, and 200 units. The Blue assaulted force is 50 units in all cases (because the assault width is fixed.) These sample calculations are shown in Figures (XIII.D-2) - (XIII.D-4). The initial and final Red assault to Blue assaulted force strengths are summarized in Table XIII.D.1.

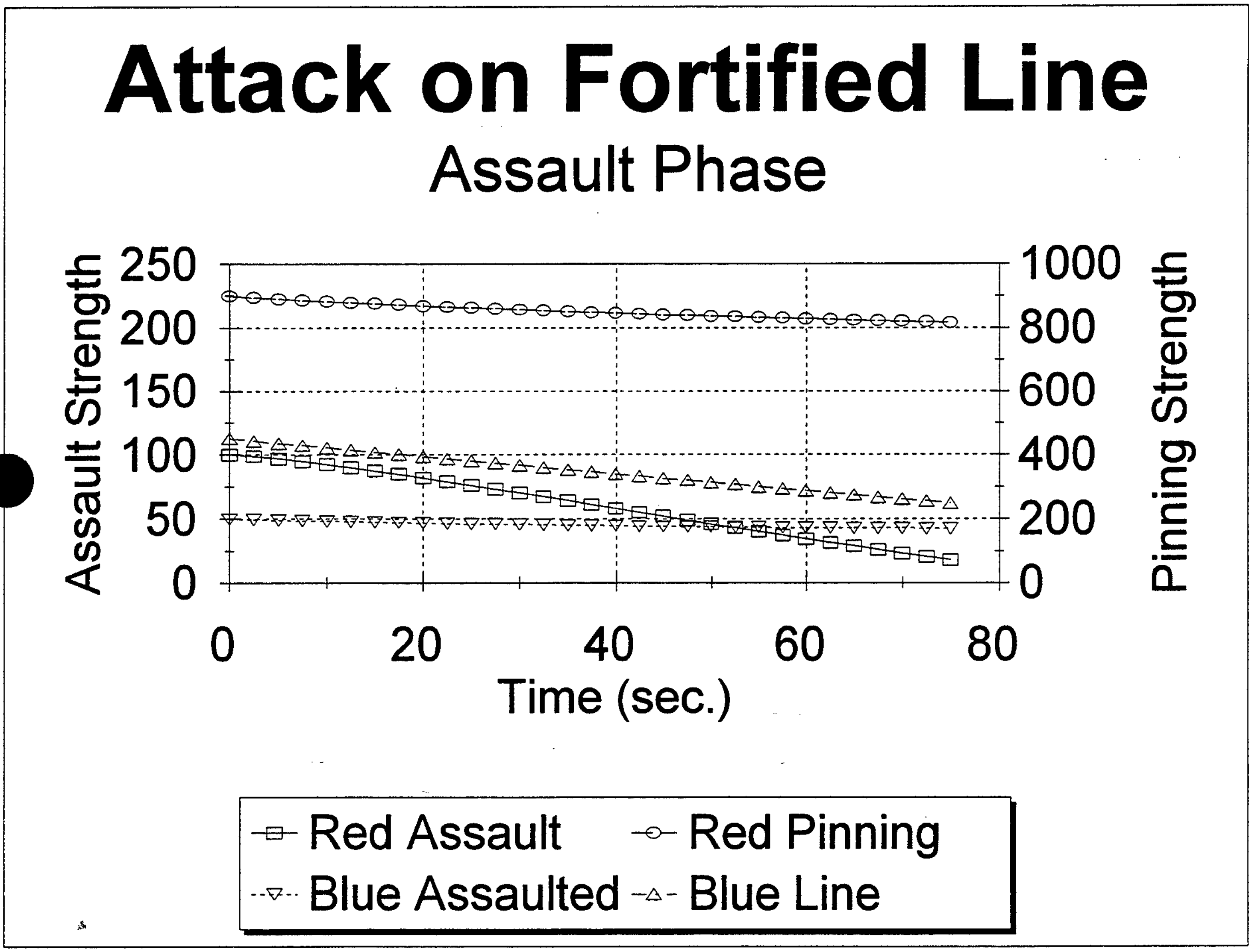


Figure XIII.D.2

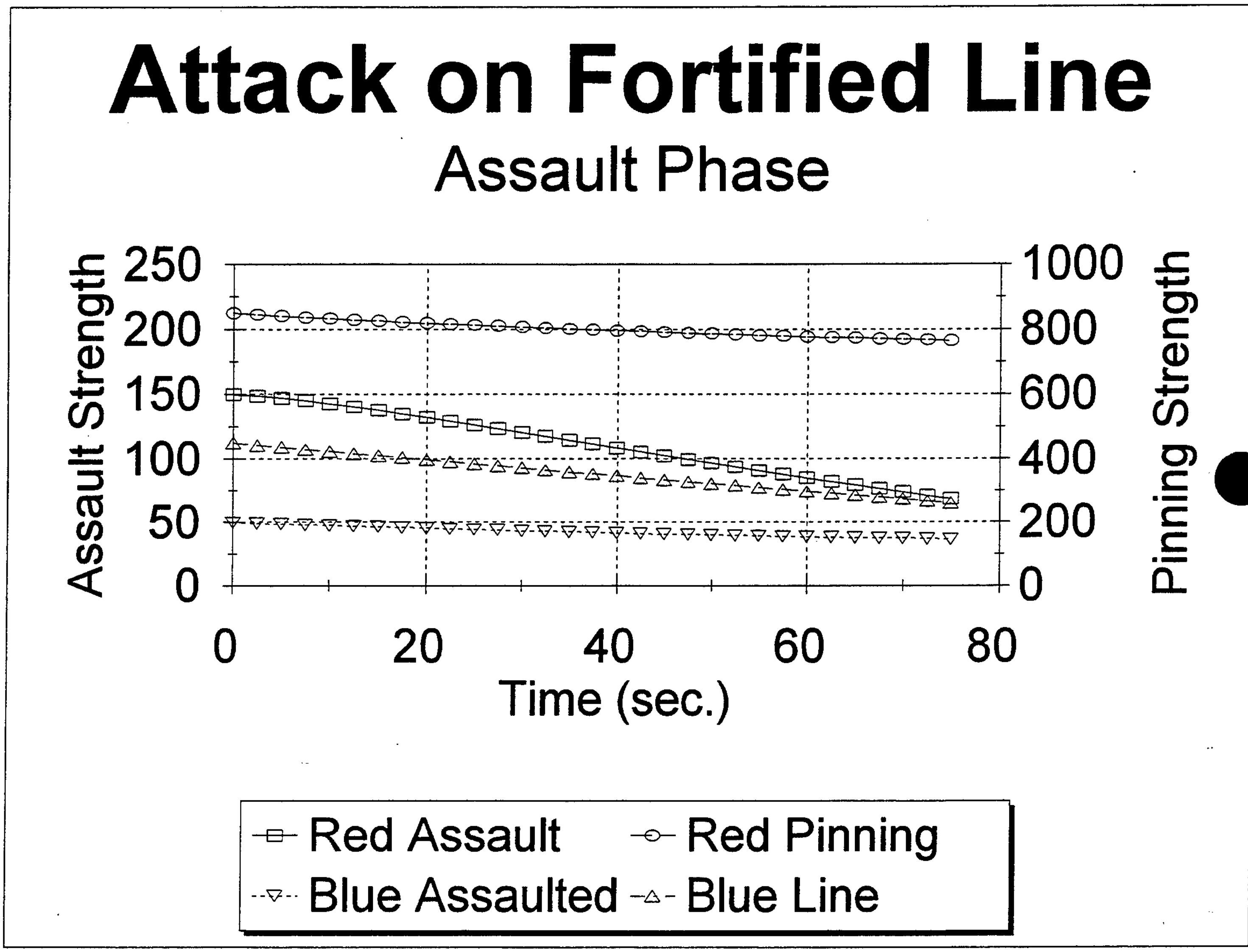


Figure XIII.D.3

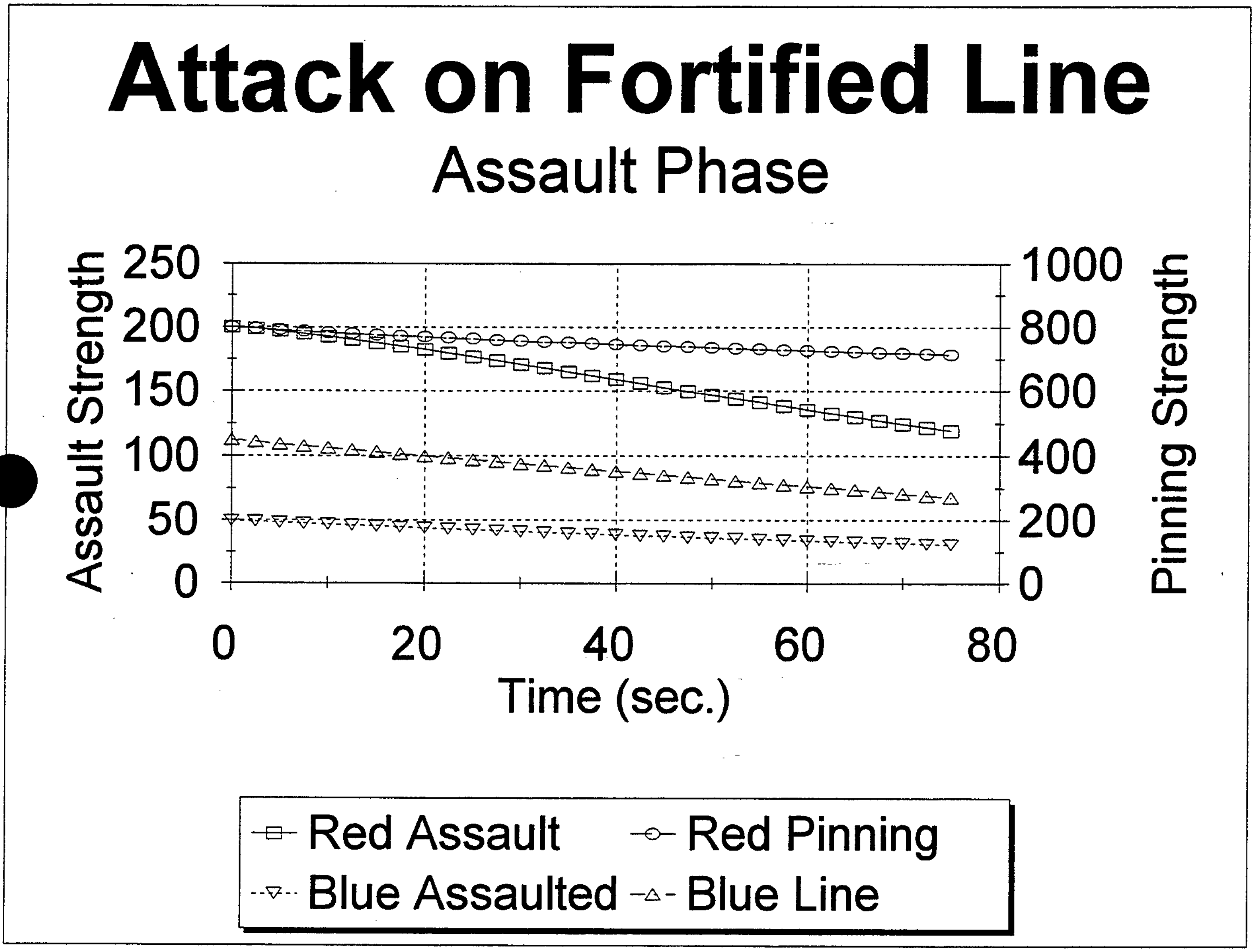


Figure XIII.D.4

Table XIII.D.2. Variable Assault Width Results

Figure #	Initial Red:Blue Force Ratio	Completed Red:Blue Force Ratio
XIII.D.5	8	26.86
XIII.D.4	4	3.86
XIII.D.6	2.67	1.85

The next set of calculations repeated the parameters for the calculation shown in Figure (XIII.D-4), Red assault force = 200 units, except to vary assault width at 50 and 150 meters. These calculations are shown in Figures (XIII.D-5) and (XIII.D-6), and the force ratios are summarized in Table XII.D.2.

We cannot directly compare the results presented in the two tables because of the effect of the peripheral Blue units engaging the Red assault unit. It is possible however to draw several conclusions from these data albeit that much of it is merely confirmatory of historical and doctrinal knowledge. First, the assault force needs to be larger than the assaulted force. Second, the assault force should be as concentrated as feasible. These are the merely confirmatory results. What is not merely confirmatory is the decidedly non-linear relationship between initial and completed force ratio. Note that in both calculations an initial force ratio of 4 gives a completed force ratio of 3.86, but that an initial force ratio of 8 gives a completed force ratio of almost 27! This non-linearity gives some indication of the degree of discussion in the tactical literature, even today, of the size of assault force necessary to force a successful breakthrough. If this function is highly non-linear, as these data indicate, and we have only limited actual, historical data, overlaid with many different values of parameters, which will change these values, then we may expect lively debate on this question.

We must strongly emphasize that these data in these two tables are highly dependent on the timing parameters that we have selected for these examples, and to a lesser extent on other parameters including the involvement of peripheral Blue units, and even on the structure of the model itself. Nonetheless, they do indicate that there is a decidedly non-linear behavior between initial and completed force ratios. General understanding of the dynamics predicted by models such as these require many more calculations. We must recall that our purpose here is to demonstrate the use of Lanchester Attrition Theory to build the model, not to exhaustively exercise the model to complete understanding. That exercise may be

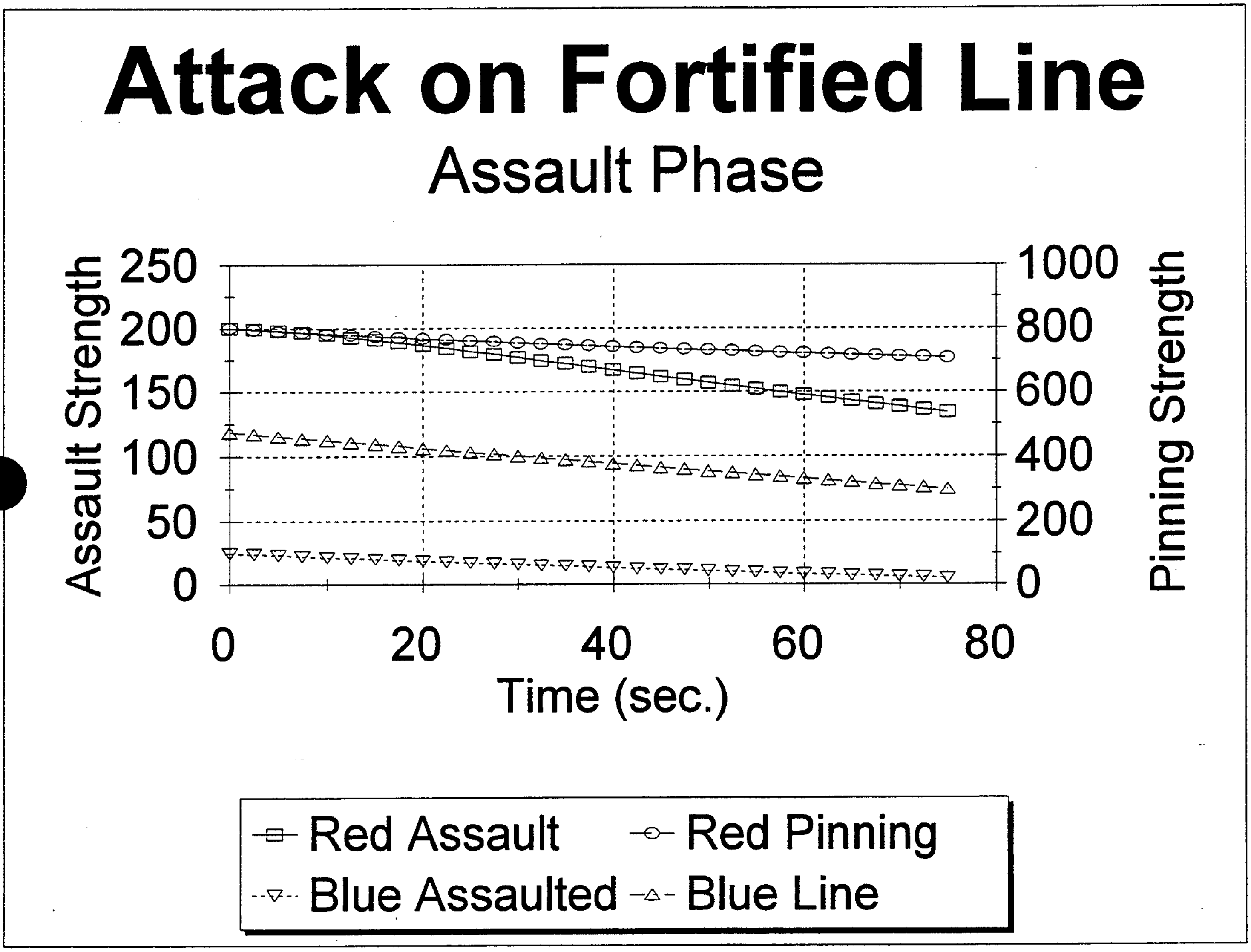


Figure XIII.D.5

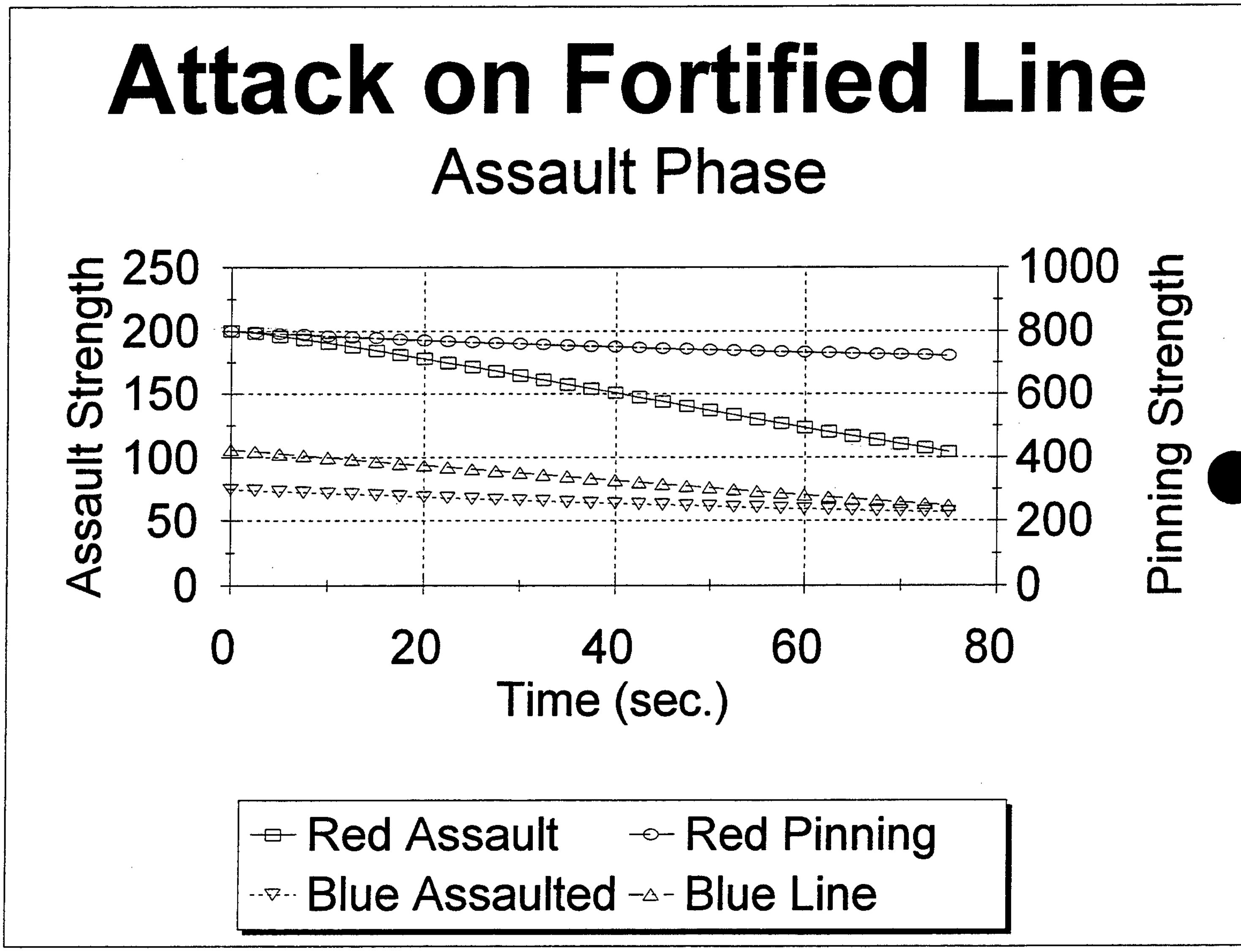


Figure XIII.D.6

pursued by the student at leisure with the understanding that other variations on the model structure should be considered.

How much is enough completed force ratio? This question can only be answered in the context of the second phase of the battle, and we must now turn to a descriptive model of this phase.

Having completed the assault phase, that is, Red units have closed the distance to the fortified line and are now within the line. At this point, the battle shifts to a mixture of what we have already modeled in the assault phase, and the meeting engagement that we modeled in the previous section. Because of the actual dynamics of the battle, this model can become very complicated very rapidly if we try to make it too accurate. To see this, let us take a snapshot of the battle at the instant that the Red assault force is inside the fortified line.

At this instant, the Red assault force, R_1, and the Blue assaulted force, B_1, are in close contact, if not intermixed, and are executing what we may think of as a normal quadratic Lanchester engagement with essentially identical attrition rates (unless Blue has prepared a secondary defensive line - an elaboration that we shall ignore but note as modelable.) The other two forces, R_2 and B_2, continue to fight as before, although probably both commanders now begin to move units of these forces to reinforce the breakthrough meeting engagement. The Red commander sends some or all of R_2 to reinforce R_1, and may leave some of R_2's units to continue to engage the remainder of the Blue units on line, or to try another assault. The Blue commander may send some of his B_2 to reinforce B_1 while leaving some units to man the line. In both cases, the reinforcing units, and in Red's case, the possible second assault, take time to arrive as reinforcements, and because they are spread along the length of the fortified line, their arrival will be spread out unless they take time to group and they can be attrited while they move.

It is easy to see that where our assault phase model had two force components for each side, the post-assault or breakthrough phase model must have at least three force components for each side, and that for good accuracy, we may have to change our representation of the forces entirely to more continuously reflect their distribution in space (and thereby time.) We can however, construct a simple three component model of this phase of the battle. This three component post-assault model consists of four differential equations per side since we must account for the fact that the breakthrough does not necessarily occur at the center of the line. The attrition differential equations for the Red force are given by

$$\begin{aligned}
\frac{dA_1}{dt} &= -\alpha_1 B_1 + a_1(t) \\
\frac{dA_2}{dt} &= -\alpha_2 B_2 \frac{A_2}{A_2 + A_{3<} + A_{3>}}, \\
\frac{dA_{3<}}{dt} &= -\alpha_2 B_2 \frac{A_{3<}}{A_2 + A_{3<} + A_{3>}} + a_{2<}(t) - a_{1<}(t), \\
\frac{dA_{3>}}{dt} &= -\alpha_2 B_2 \frac{A_{3>}}{A_2 + A_{3<} + A_{3>}} + a_{2>}(t) - a_{1>}(t),
\end{aligned} \qquad \text{(XIII.D-6)}$$

and for the Blue force by

$$\begin{aligned}
\frac{dB_1}{dt} &= -\beta_1 A_1 + b_1(t) \\
\frac{dB_2}{dt} &= -\beta_2 A_2 \frac{B_2}{B_2 + B_{3<} + B_{3>}}, \\
\frac{dB_{3<}}{dt} &= -\beta_2 A_2 \frac{B_{3<}}{B_2 + B_{3<} + B_{3>}} - b_{1<}(t), \\
\frac{dB_{3>}}{dt} &= -\beta_2 A_2 \frac{B_{3>}}{B_2 + B_{3<} + B_{3>}} - b_{1>}(t),
\end{aligned} \qquad \text{(XIII.D-7)}$$

where:

$$\begin{aligned}
a_1(t) &= a_{1<}(t) + a_{1>}(t), \\
b_1(t) &= b_{1<}(t) + b_{1>}(t).
\end{aligned} \qquad \text{(XIII.D-8)}$$

These equations are similar in structure to those for the meeting engagement model, equations (XIII.C-1) and (XIII.C-2), in that there are reinforcement terms. Let us note the similarities and differences in equations (XIII.D-6) and (XIII.D-7). First, the A_1 and B_1 attrition differential equations, which are identical in form to equations (XIII.C-1) and (XIII.C-2), describe the engagement between the two components at the breakthrough. The A_2 and B_2 force components are the forces left to engage along the line. These forces, respectively, engage both the B_2 and B_3, and A_2 and A_3 components. Note that their fire is equally distributed over both components (the fractions in the last three equations in equations (XIII.D-6) and (XIII.D-7).)

A major difference in these equations is the presence in the A_3 equations of reinforcement terms a_2 that are not present in the B_3 equations. This is a direct result of the geometry of the problem and will be clearly expressed mathematically below. As we stated earlier, we assume that each commander takes a fraction of his second component force to reinforce the forces in the breakthrough area. For the Blue force, these reinforcements merely have to travel to that area while for the Red force, these reinforcements must first travel behind their line to the break in their line and then down to the breakthrough area. Thus, relatively speaking, the Red reinforcements must travel a distance r_0 further than the equivalent Blue units. Since they are assumed to be traveling behind their line, they are not engagable. The arrival of these units to travel from their line to the fortified line gives rise to the a_2 reinforcement terms.

Before we define the reinforcement terms and the boundary conditions, it is necessary to define some parameters and to introduce a bit of notational mathematics. First, per our assumptions, each commander assigns part of his second (assault) component force (A_2 or B_2) to be reinforcements. Designate these parts as fractions r_A and r_B, respectively, and stipulate that they move at speeds s_A and s_B, respectively. Designate the center of the breakthrough with the coordinate x with the provisos that $x \geq w/2$, and $x \leq L - w/2$. (This merely assures that the breakthrough width lies within the line.) Next, we must define the step function $\eta(z)$ such that

$$\begin{aligned} \eta(z) &= 1, z > 0, \\ &= 0, z < 0. \end{aligned} \qquad \text{(XIII.D-9)}$$

We may now define boundary conditions and reinforcement rates.

Recall that we shift models from assault phase to post-assault phase at $t_{assault}$. For our model, we do not add in a delay time for breaching the fortified line, although this is possible. Since our primary interest at the moment is in terms of Civil War filed fortifications, we assume that they impose no exceptional impediment to movement that would substantially delay the assaulting force. For consideration of more elaborate (and impedimential) fortifications, a delay time may be added to $t_{assault}$. Further, recall that we have assumed that (outside the breakthrough area,) the forces are uniformly distributed over the lines.

With these factors in mind, we may now establish the boundary conditions on equations (XIII.D-6) and (XIII.D-7). Since these equations become valid for times greater than $t_{assault}$, we introduce a new time variable t' defined by $t' \equiv t - t_{assault}$ to simplify notation and calculation. Note that t' is <u>not defined</u> for value less than 0 and that $dt' = dt$, so that our differential equations keep the same form in the new coordinate system. With all this in hand, we may write boundary conditions as

POST-ASSAULT PHASE ASSAULT PHASE

$$
\begin{aligned}
A_1(0) &= A_1(0),\\
B_1(0) &= B_1(0),\\
A_2(0) &= (1 - r_A)\, A_2(0),\\
B_2(0) &= (1 - r_B)\, B_2(0),\\
A_{3<}(0) &= 0,\\
B_{3<}(0) &= r_B\, B_2(0)\, \frac{x - \frac{w}{2}}{L - w},\\
A_{3>}(0) &= 0,\\
B_{3>}(0) &= r_B\, B_2(0)\, \frac{L - x - \frac{w}{2}}{L - w}.
\end{aligned}
\qquad \text{(XIII.D-10)}
$$

Note that the quantities on the left-hand-side are assault phase variable values at t' = 0, that is $t = t_{assault}$, while the quantities on the right hand side are post-assault phase variable values at the same time - boundary conditions. In particular, the forces at the breakthrough area are the same, while the forces on line, for both sides, are reduced by fractions r_A and r_B. The Blue reinforcing forces are created from those forces removed from the Blue line, and are apportioned into two sub-components indicated by the symbols < and > to indicate whether they are below or above the breakthrough area (i.e. to Blue's left or right in terms of military usage.) Note also that the Red reinforcing sub-components are initially zero since the Red forces drawn from the line must reach the assault corridor to be counted - these forces will explicitly appear in the a_2 reinforcement rates - they have not magically been lost!

We mentioned earlier that this use of homogeneous aggregation of forces was an approximation, and that we should really use a continuous spatial representation of the forces to account for the movement and attrition of our forces modeled here, particularly the reinforcing ones. Now that we are finally able to write out the form of the reinforcement rates, we can make note of the substantial nature of that approximation. While we can account for the differential travel times of the reinforcing forces, we must lump those forces together for attrition purposes. This may not be very realistic as we can conceive that these forces would take progressively greater losses as they get nearer the fortified line (for the Red forces,) or nearer the breakthrough area (for both forces.) Notice also that the reinforcing forces do not fight - they only move - until they become part of the reinforcement rates.

The reinforcement rates for the breakthrough area forces may be given simply as

$$a_{1<}(t') = A_{3<}(t') \frac{s_A}{\rho + x - \frac{w}{2} - s_A t'} \eta\left(t' - \frac{\rho}{s_A}\right) \eta\left(\frac{\rho + x - \frac{w}{2}}{s_A} - t'\right),$$

$$a_{1>}(t') = A_{3>}(t') \frac{s_A}{\rho + L - x - \frac{w}{2} - s_A t'} \eta\left(t' - \frac{\rho}{s_A}\right) \eta\left(\frac{\rho + L - x - \frac{w}{2}}{s_A} - t'\right),$$

$$b_{1<}(t') = B_{3<}(t') \frac{s_B}{x - \frac{w}{2} - s_A t'} \eta\left(\frac{x - \frac{w}{2}}{s_B} - t'\right),$$

$$b_{1>}(t') = B_{3>}(t') \frac{s_B}{L - x - \frac{w}{2} - s_A t'} \eta\left(\frac{L - x - \frac{w}{2}}{s_B} - t'\right). \qquad \text{(XIII.D-11)}$$

If we examine the left-hand-sides of these equations, we may see the logical structure of the reinforcement rates. The first terms are the current force strengths of the reinforcing forces. The next terms are the arrival rates of the portions of these forces. If we take each of these fractions and multiply each by a time increment Δt, then this result is the fraction of the reinforcing force that arrives at the breakthrough area in that time increment. The denominators represent the linear form of the reinforcing forces - the left (relative to Blue,) or < forces are initially a line x - w/2 long, while the right or > forces are initially a line L - x - w/2 long. Because these lines are moving, and once they reach the breakthrough area are getting shorter, we must reduce the length of that line. The addition of ρ in the denominator for the Red reinforcement rates accommodates the additional time required for them to arrive at the breakthrough area - we could equally well have inserted a delay time subtracted from t' equal to $t_{assault}$ that would have the same effect.

The first η functions in the Red reinforcement rates reflect the time required for the first units in the reinforcing forces to traverse the assault corridor between the lines. The η functions in the Blue reinforcement rates and the second η functions in the Red reinforcement rates reflect the time for the last units in the reinforcing forces to reach the breakthrough area. Note that the definition of the η function means that these functions initially have a value of one but become zero when time progresses enough to make the arguments negative. Thus the Red reinforcement rates are zero until time progresses to make the argument of the first η function positive, and continue until time progresses to make the argument of the second η function negative - the product of the two η functions is effectively a unit height function that has zero value before and after reinforcements arrive, and one during. The Blue reinforcement rates occur initially, because of the shorter travel distance, but cease later, so only one η function is required.

This completes the post-assault phase model. Like the assault phase model, this model does not obviously possess a state solution, and indeed, because it has so many time dependent terms, we do not expect it to. Unlike that model, I do not believe that this model has a simple analytic solution largely because of the way we have implemented the attrition of the line and reinforcing forces. I cannot categorically say that such a solution does not exist, merely that I cannot see an obvious approach and have not tried to solve the equations. Regardless of this, we can calculate numeric solutions using the same techniques that we have practiced before, and we shall now proceed to examine several examples.

For our examples, we extend the calculations previously presented in Figures (XIII.D.4) - (XIII.D.6) for the assault phase to the post-assault phase. These calculations are shown in Figures (XIII.D.7) - (XIII.D.9), respectively. We have used movement speeds that are the same as used in the assault phase, now for both forces, recognizing that these speeds may be somewhat high for the Civil War era. For reinforcement fractions, we have used 50% for both sides, and have assigned the breakthrough area to be the mid-point of the line (i.e. $x = L/2$.)

If we examine Figure (XIII.D.7), we see that despite the delay in Red receiving reinforcements, the battle seems to clearly be in Red's favor. Indeed, the Blue line is completely attrited - our Red on Blue line attrition rate may be too high for historical accuracy - we have seen this trend in the calculations for the assault phase. Once the Blue line is attritted, the Red line has nothing to do, since the Blue reinforcing force has gone away - fully deployed in the breakthrough area, and the Red commander could deploy some of that force (about 300 units of the original 1000) as reinforcements. This is probably not necessary as by $t' = 200$ sec., the Red breakthrough force has a strength of about 400 while the Blue force has been reduced to about 50 - a force ratio of about 8! This can be compared to the force ratio at the end of the assault phase of 3.86. Note also that at $t' = 75$ sec., when Red reinforcements begin to arrive, that the force ratio in the breakthrough area is about 1.

Figure (XIII.D.8) displays similar behavior except that because the assault width has been reduced to 50 m., there are fewer Blue units in the area, and Blue reinforcements have further to travel. The differences between this case and the previous one are only marginal.

Figure (XIII.D.9) shows a more complex case, but substantially the same result. Admittedly, Blue has a favorable force ratio for a longer period of time, from about $t' = 60$ sec., to $t' = 80$ sec., but by $t' = 200$ sec., Red has still achieved a force ratio of better than 5, and the Blue line has been completely attrited.

Before concluding this section, some reflection on these models is in order. Clearly, we could have constructed them in other ways to reflect other tactical decisions or doctrinal rules. That we have not done so does not imply that this model is inherently the correct one for all situations. Our intent was to demonstrate that models for complex tactical situations can be readily constructed using Lanchester Attrition Theory and equally readily examined. The student should now be able to construct different models reflecting other tactical and doctrinal considerations and exercise them for insight.

For this pair of models in particular, several key points can be made. They clearly demonstrate the unrealness of the conclusion condition of Lanchester theory as a criterion for victory. The need for a battle termination model such as Weiss' in the previous chapter is obvious. From an historical standpoint, our estimation of attrition rates here, based on our limited coverage of this area of theory, is equally clearly in error. In particular, the attrition rates against the units in the fortified line are too high. Were we to adjust these rates downward, we would expect some change in the post-assault phase results.

It may be argued that the movement rates (speeds) used here are also too large. To address this question, we repeated the calculations presented in Figures (XIII.D.6) and (XIII.D.9) using halved movement rates - 2 m/sec. These calculations are shown in Figures (XIII.D.10) and (XIII.D.11), where we have adjusted the durations of the two phases proportionally. Aside from the increased attrition to all forces due to the longer times, there is no substantive change in the shapes of the curves.

On the positive side, the two models do confirm the observations of Weiss. If a substantial assaulting force can reach the fortified line and create a break-through, it is difficult for the defending force to eject them. Admittedly, in our calculations, we never investigated the case where the assaulting force did not satisfy this condition, primarily because of the tactical question of what to do about the Red (assaulting) reinforcing force.

In conclusion, let me reiterate that the purpose of this modeling section was primarily to demonstrate how Lanchester Attrition Theory can be used to rapidly and simply develop fairly complex tactical situation models that can then be easily exercised (using spreadsheets) to gather insight. The professional soldier may be quick to decry the fallacies and limitations of this particular model, either because it does not accurately reflect some tactical or doctrinal wisdom (e.g., either don't allow the assault force to fire, or reduce their movement rate to reflect firing,) or because it does not reflect the situation accurately enough. Both of these objections can be accommodated by changing parameter values or introducing new complexity into the model. The student should now be capable of taking exactly these actions,

but I counsel caution - inject new complexity only after you understand the model at each level.

XIII.E. Survival of Battle

With this section, we now return to consideration of the historical data. In particular, we now consider the historical evidence, provided by our data bases, of the statistics associated with what average per-centum losses have been. This completes the effort begun in the previous chapter where we presented these types of calculations for our Civil War data base. Now, we perform these same types of calculations with our other four data bases.

As a reminder, the calculation that we perform is to take the initial, final force strength data in our data bases, and perform a linear regression on this data to an equation of the form,

$$S_{final} = \sigma S_{initial}, \tag{XIII.E-1}$$

where: $S_{initial}$, S_{final} are the force strengths, and σ is the slope of this line. As we noted earlier, the form of this straight line; that is, zero intercept, is predicated on the idea that an initial force of zero strength must result in a final force of zero strength; this is just fancy mathematical talk for what is really common sense. The selection of a straight line is based on our visual inspection of the cross plots of the force strengths presented in a previous chapter. As we noted then, there may be additional structure, indicating either higher order effects, or a break in the data trends, but for our purposes here, to examine the values of σ, which represents the average survival fraction, the straight line is our choice.

Now. we may examine these regressions. For the individual data bases, (except the Civil War data base, presented in the preceding chapter,) these results are presented in Table XIII.E.1. The amazing result here is the consistency among the first three data sets, all having σ value of approximately 0.84! All three of these data sets also have comparable levels of error in the fit, of the order of 1%, and all four data sets have very high correlations, greater than 0.96. The exception is the data for World War I, which has a considerably smaller σ value, and a larger error (but the smallest number of battles by a factor of three.)

This set of regressions indicates that over a very great period of time, since the Normal (or Brassey's) and Short Battle data sets span a very long period of time, the average losses in battle tend to be about 16% per battle. This indicates that battles with very large or very small losses tend to be rare. The World War I results tell us what we already know, that these battles were much more bloody

than the norm by about 50% (in per-centum losses.) The small values of the errors and the high values of correlation give us a high degree of confidence in this.

Table XII.E.1. Force Strength Linear Regressions of Remaining Data Sets

Data Set	# of Battles	σ	Stan-dard Error	R^2
Normal	215	0.838	0.009	0.962
Osipov's	76	0.845	0.01	0.966
Short	144	0.845	0.007	0.985

Table XIII.E.2. Winner-Loser Force Strength Regression Results

Data Set	# of Battles	σ	Standard Error	R^2
Osipov's Winners	37	0.872	0.013	0.969
WW I Winners	12	0.794	0.022	0.978
Osipov's Losers	37	0.808	0.013	0.969
WW I Losers	12	0.765	0.023	0.971

Three of these data sets provide us with additional break-down on the data. In particular, the Osipov and World War I data sets identify Winners and Losers, and the Short and World War I data sets identify Attackers and Defenders. We may also subject these data set divisions to the same type of regression. The results of Winners and Losers is presented in Table XIII.F.2, and for Attackers and Defenders in Table XIII.E.3.

We continue to see very good correlations with these regressions, although the errors are considerable larger than before, an expected result given the division of the data sets in half in terms of the number of data points. That is, for small errors (less than 10% say,) we would expect the error to double if we decrease the size of the data set by a factor of two. Nonetheless, the data are still good enough that we may draw some fairly obvious conclusions. Recalling that Osipov's Battles are essentially Nineteenth Century battles (actually Napoleonic to just before World War I,) and that the World War I Battles, which follow directly, chronologically, after Osipov's, are particularly bloody, we may observe that in general, winners take fewer losses than losers, 13% versus 19% for Osipov's, and 21% versus 23% for World War I, and that attackers take fewer loses than defenders, 16% versus 18% for Osipov's, and 21% versus 23% for World War I.

Table XIII.E.3. Attacker-Defender Force Strength Regression Results

Data Set	# of Battles	σ	Standard Error	R^2
Short Attackers	72	0.845	0.007	0.988
WW I Attackers	12	0.778	0.024	0.973
Short Defenders	72	0.823	0.010	0.978
WW I Defenders	12	0.765	0.024	0.964

Clearly, this trend for defenders is contrary to what is commonly considered to be conventional military wisdom, but it tends to confirm the claims made that the advance of technology reduces the advantage of the defender. Certainly, we can

Table XIII.E.4. Exponential Cumulative Fractional Exchange Ratio Distributions Fits

Data Set	Function	Fitted Parameter	D Value
Osipov	Exponential	1.0243	0.3693
	Gaussian	1.093	0.1986
World War I	Exponential	0.6130	0.3217
	Gaussian	0.3722	0.1424

visualize that as weapons become longer range and more lethal, and attackers can carry armored protection with them, the value of prepared defenses to shield the defender and limit the concentration of the attacker becomes reduced.

Table XIII.E.5. Inverse Polynomial Cumulative Fractional Exchange Ratio Distribution Fits

Data Set	Polynomial Order	D Value
Osipov	1	0.3384
	2	0.2519
	3	0.2255
World War I	1	0.3878
	2	0.2863
	3	0.2026

Finally, we may examine the statistics of the distributions of these data using the cumulative Kolomogorov-Schmirnov technique described in the appendi ces. In particular, we may examine the probability of willing, ala Weiss, in terms of the Fractional Exchange Ratio distribution for the Osipov and World War I data

Table XIII.E.6. Fractional Exchange Ratio Distribution Fits for Brassey's Battles

Data Period	Exponential Fit	D Value	Gaussian Fit	D Value
Pre-Gunpowder	3.342	0.2433	8.0280	0.2976
Early Gunpowder	4.2212	0.1223	14.9994	0.1339
18th Century	5.8783	0.1197	21.4883	0.2661
1800-1825	6.2151	0.1189	22.6746	0.2212
1826-1860	4.0326	0.1981	7.6335	0.4964
1861-1876	8.2125	0.2075	28.3421	0.3588
1876-1899	5.0204	0.0892	15.5878	0.3294
20th Century	4.1855	0.2420	16.8580	0.1261

sets. In both sets, we find that the distribution is better fit by a gaussian rather than the hypergaussian or the inverse polynomial, although the difference is small. Of particular note here is the F_{ER} functional relationship does not seem to support the unequivocal use of the third power. We present these data in Tables XIII.E.4 and XIII.E.5.

Table XIII.E.7. Fractional Loss Distribution Fits for World War I Battles

Segment	Exponential Fit	D Value	Gaussian Fit	D Value
All	3.3736	0.2548	10.4845	0.1156
Attacker	3.7852	0.2769	15.1318	0.1066
Defender	2.6063	0.3356	7.0235	0.1794
Winner	3.3132	0.2087	9.8394	0.1574
Loser	2.8898	0.3807	9.3604	0.2700

For the Brassey's Battles data set, we may examine the F_{ER}, not as a measure of probability of winning, but as a simple indicator of distribution. The fitting and error data for these are presented in Table XIII.E.6. Two things may be noted from this table. First, the value of the fitted parameters vary considerably by period, which we must hasten to note are somewhat arbitrary. In particular, the periods of the Napoleonic and American Civil Wars are particularly described by small F_{ER}s (large fit parameters.) Second, except for the Twentieth Century, the fits are better for the exponential function than for the Gaussian.

If we now turn to examine fractional losses (rather than F_{ER}s,) in the same manner, then we find somewhat different results. For the World War I data set, these are summarized in Table XIII.E.7. The same data for Osipov's data set are summarized in Table XIII.E.8. While the World War I battles seem to more clearly be fit by gaussian distributions, the Osipov data are not as clear cut.

This concludes our discussion of historical data for now. If there is a lesson to be learned here, it seems to be that the lack of detail available precludes any simple and concise approach to either supporting or denying Lanchester theory. There are clearly trends and analyses that may support both sides of this argument, but

either side is sufficiently strong to be conclusive. We must therefore proceed in the sure uncertainty that Lanchester theory may be used as an indicative but not as a definitive or even predictive tool of war.

Table XIII.E.8. Fractional Loss Distribution Fits for Osipov's Battles

Segment	Exponential Fit	D Value	Gaussian Fit	D Value
All	5.3561	0.2407	21.3221	0.1340
Stronger	6.3580	0.1758	27.7845	0.1892
Weaker	4.3706	0.3057	16.3954	0.1043
Winner	6.5535	0.1594	28.9819	0.2009
Loser	4.2827	0.3523	16.0052	0.1530

References

1. Weiss, Herbert K., "Combat Models and Historical Data: The U. S. Civil War", Operations Research, 14 (5), September-October 1966, pp. 759-790.

2. Carnahan, Brice, H. A. Luther, and James O. Wilkes, **Applied Numerical Methods**, John Wiley and Sons, Inc., New York, 1969, p. 71.

3. McWhiney, Grady, and Perry D. Jamieson, Attack and Die: Civil War Military Tactics and the Southern Heritage, The University of Alabama Press, University, AL, 1982, pp. 48-58.

4. Griffith, Paddy, **Battle Tactics of the Civil War**, Yale University Press, New Haven, 1987, pp. 73-90.

XIV. CONCLUSIVE BATTLES[a]

XIV.A. INTRODUCTION

In this chapter, we finally come to treat with that special class of battles first introduced by Lanchester - conclusive battles.[1] Of course, these battles are the special case where the combat is carried to the point where one side is completely attrited - it has no units left, thus the term conclusion. In terms of the state solution of the attrition differential equations for general attrition order n,

$$\Delta_n \equiv \alpha B_0^{\,n} - \beta A_0^{\,n}, \qquad \text{(XIV.A-1)}$$

this situation is predictable in magnitude and side by the magnitude and sign of the Δ_n. If $\Delta_n > 0$, then the Blue force survives, and has a conclusion force strength of

$$B_c = \left[\frac{\alpha B_0^{\,n} - \beta A_0^{\,n}}{\alpha}\right]^{\frac{1}{n}}, \qquad \text{(XIV.A-2)}$$

while if $\Delta_n < 0$, the Red (Amber) force survives, and has a conclusion force strength of

$$A_c = \left[\frac{-\alpha B_0^{\,n} + \beta A_0^{\,n}}{\beta}\right]^{\frac{1}{n}}. \qquad \text{(XIV.A-3)}$$

Of course, if $\Delta_n = 0$, then the special case of a conclusive draw occurs where both force strengths are attritted to zero units.

XIV.B. Historical Data

A question which naturally arises is "How often do conclusive battles occur in history?" As a first step in this analysis, we reviewed three readily available dictionaries of battles: Dupuy and Dupuy's **The Encyclopedia of Military History**,[2] Eggenberger's **An Encyclopedia of Battles**,[3] and Laffin's **Brassey's Battles**.[4] While Dupuy and Dupuy do not state how many battles they include, Eggenberger claims to cover 1,560 battles and Laffin more than 7,000. Even with overlap, a

[a] This chapter is based on a paper of the same title written in 1992 by CPT Lawrence Phillips and myself, and submitted for publication to the journal, Military Review. It was returned by the journal without comment as unacceptable. This is a concrete example of the difficulty of publishing Lanchester Theory or military analytical papers; they are frequently too old hat for the OR journals, and either too analytical or politically unacceptable for the military journals.

conservative estimate is that the three volumes chronicle at least 10,000 battles. These are three of the major sources of data that we have already used in the development of this book.

We cannot claim that these three sources constitute an exhaustive survey of battles, or even less that our search of them has been exhaustive. We do believe, however, that the list of battles we present below is representative of conclusive battles. In our search for conclusive battles, we have limited our acceptance to those battles for which some numeric force strengths were provided." The data provided by the authors of these three works has been taken as authoritative. An extensive examination of more detailed sources was beyond the scope of this effort. The 10 battles that we were able to identify as conclusive are summarized in Table XIV.B.1 and described below.

The battles are arranged in chronological order. While they span a period from 480 B.C. (Thermopylae 1) through 1879 (Isandhlwanda) half (5 of 10) are nineteenth century. The initial inclination is to attribute this to better record keeping, although upon analysis, a second contributing factor emerges. Many of these battles are famous, far beyond what might be expected from the force strengths involved. Of the 10, 6 are battles which have become mythic within different cultural groups. We review each battle briefly.

Thermopylae 1: In this battle, the Greeks were defending the Thermopylae Pass against the invading Persian forces under Darius I (son of Xerxes) to prevent entry into Attica proper. After a three day defensive battle, the Persians found an alternate pass through the mountains and successfully flanked the Greeks. The part of the battle reported here is the successful rear guard holding action of the Spartan contingent under "King" Leonidas 1. The Persian campaign was subsequently terminated by the Greek victory at the naval battle of Salamis. This battle is considered mythic because of the courage and dedication of the Spartan force.

Teutoburger Wald: In this battle, a larger force under the Chief of the Cherusci, Arminius, surprised a Roman force under Publius Quintillus Varus in the Teutoburger Wald (Forest) as they marched to crush a German rebellion. Over a three day battle, the Roman forces were destroyed. This battle demonstrates the superiority of irregular, light forces over heavy infantry and cavalry in irregular, close terrain. Although not mythic, this battle is historically significant because it fixed the northern boundary of the Roman Empire at the Rhine River.

Table XIV.B.1 Statistics on Conclusive Battles

Battle	Date	Winner	Initial Strength	Final Strength	Loser	Initial Strength	Final Strength
Thermopylae I	-400	Persians	100,000	-	Spartans	300	0
Teutoburger Wald	9	Cherusci	-	-	Romans	20,000	0
Antioch II	1119	Turks	-	-	Crusaders	-	-
Mohi	1241	Mongols	-	-	Poles	100,000	70,000
Bannockburn	1314	Scots	8,000	4,000	English	15,000	few
Alamo	1836	Mexicans	3,000	1,400	Texans	188	0
Camerone	1863	Mexicans	2,000	-	French	65	3
Massacre Hill	1866	Sioux	2,000	1,940	U.S.	81	0
Little Bighorn River	1876	Sioux	3,000	-	U.S.	211	0
Isandhlwanda	1879	Zulus	20,000	-	British	1,800	5

Antioch II: In this battle, a Turkish force under Ilghazi, the ruler of Aleppo, surrounded a Norman force under Roger of Salerno, the Regent of Antioch, during the night before the battle. Unable to break the Turkish line, the Normans were destroyed. The battle had little strategic significance as Ilghazi failed to press his initiative following the battle.

Mohi: In this battle, Mongol forces under Subodai made rear, flank, and frontal attacks against the Polish forces (in battle line) under King Bela IV. Although not technically a conclusive battle since approximately 30 percent of the Polish force survived, the Polish Army was routed and destroyed as a fighting force. The Mongols were subsequently unable to maintain the initiative gained in this battle because they were re-called to participate in the selection of a new Kahn following the death of Ogadai.

Bannockburn: In this battle, Scottish forces under King Robert I (the Bruce) initially defended uphill against an assault by an approximately twice larger and heavier English force under King Edward II. After initially repulsing the English forces, the Scots seized the initiative by counterattacking across boggy ground where the lighter Scottish forces destroyed the heavier English forces. This battle is mythic among Scots since it secured Scottish independence. (Edward was unable to mount any further invasions because of civic and Irish unrest.)

Alamo: A small force of Texan revolutionaries occupied and fortified a

Franciscan mission, the Alamo, which lay astride the supply and communication lines of the Mexican forces commanded by Antonio de Santa Anna. The Texan intent was to delay Santa Anna's progress north to buy time for the main Texan forces to train and equip. With this threat to his lines of supply and communication, Santa Anna was forced to lay siege to the Alamo, delaying him for 12 days. The Texan defenders were destroyed at great price to the Mexicans. Following this, Santa Anna fought seven more battles with the Texan revolutionaries before being decisively defeated at San Jacinto River. Both battles are mythic among Texans.

Camerone: In this battle, a small French Foreign Legion force held off a much larger Mexican revolutionary force for 10 hours. The battle is mythic within the French Foreign Legion as an example of Legionnaire devotion and courage.

Massacre Hill: In this battle, a small force of U.S. Cavalry counterattacked a much larger force of Sioux warriors under Crazy Horse and Red Cloud to relieve a wagon train. Apparently, the U.S. force was unaware of the size of the Sioux force when they counterattacked. Although the counterattacking force was destroyed, the Sioux were unable to follow up on their initiative.

Little Bighorn River: In this battle, 12 troops of the U.S. Seventh Cavalry under George Armstrong Custer attacked an encamped Sioux and Cheyenne force led by Sitting Bull, Crazy Horse, and Gall. Custer divided his force, personally commanding a force of five troops (this is the part of the battle considered here). Custer was surrounded and over-whelmed by superior numbers. This battle is mythic among Americans (especially in the U.S. Army).

Isandhlwanda: In this battle, a large force of King Cetewayo's Zulus over-whelmed a British regiment in broken terrain at the Great Rock of Isandhlwanda. This battle is mythic to the British because it represents the first defeat of British troops by "natives" in battle. The Zulu initiative was blunted almost immediately at the battle of Rorke's Drift (another mythic battle), and the Zulus were destroyed as a fighting force six months later at the Battle of Ulundi.

Wounded Knee Creek: An eleventh battle is also technically conclusive. Following the death of Sitting Bull, a Sioux group fled the reservation into the badlands of South Dakota. This group attacked elements of the Seventh Cavalry who were attempting to disarm the Sioux and return them to the reservation. In the subsequent action and return to the reservation, the Indians were destroyed. This battle enjoys mythic proportions among American Indians as a massacre. We do not include this battle in our analysis for this reason.

XIV.C. Analysis

Ten battles are not a very large sample out of 10,000, being about 0.1 percent

of that number. If this number is indicative, however, the answer to our question "How often do conclusive battles occur in history?" is "Not very often."

If we examine these 10 battles, some interesting trends emerge. Most of the battles clearly and immediately fall into two categories: First, actions where a small force defends against a larger force (hold and die;) and, second, engagements between two forces which are technologically and culturally quite disparate. The first category (hold and die battles), includes Thermopylae I, Alamo, and Camerone. The second category (cultural.disparity battles) includes Massacre Hill, Little Bighorn River, and Isandhlwanda. All of the disparity battles took place during the latter half of the nineteenth century when the British (and most European nations) primarily fought colonial wars and Americans were consolidating their continental nation.

The fact that 3 of the 10 conclusive battles are hold and die battles does not seem surprising. Battles of this type must end in conclusion unless the defending force is relieved. The.fact that all three of these battles are mythic would seem to indicate the rarity of this type of battle.

Of the three cultural disparity battles, all were losses for the more technolo gically and culturally developed force." Both of the American losses (Massacre Hill and Little Bighorn River) were the result of tactical mistakes on the part of the loser (inadequate intelligence in both cases, splitting the force in the second). The British loss (Isandhlwanda) is not so clear, although similar reasons have been offered for this loss as well (inadequate intelligence, possibly inadequate tactical resupply). It seems reasonable, therefore, to posit that a significant cause of these losses was tactical overconfidence. If so, this is a dire lesson to be learned by contemporary military forces in the United States and Europe as the likelihood of Third World involvement increases.

Of further note in these three battles of the second category, all were fought by forces on one side seeking conquest and on the other side by forces resisting conquest. This conquest seeking/resisting theme is common to both the first and second categories. Further, it must be considered that both of the American battles in this category were with a foe with an essentially common command structure (Crazy Horse and Sitting Bull). The fate of Custer and his five troops of the Seventh Cavalry also cannot be discounted - soldiers have long memories.

The four remaining battles, Teutoburger Wald, Antioch II, Mohi, and Bannockburn, have parallels with the battles of the second category. In all four battles, there is cultural, if not technical, disparity: Romans and Germanic Tribesmen; Normans and Turks; Poles and Mongols; and English and Scots.

We might argue that all of these battles represent conflicts between main stream Western (European) cultural forces (Romans, Normans, Poles, and English)

and non-main stream Western, or non-Western cultural forces (Germanic tribes, Turks, Mongols, and Scots.) All were conquest-type battles. Three (Teutoburger Wald, Antioch II, and Mohi) are characterized by inadequate intelligence on the part of the loser. Two are characterized by the use of inappropriate types of units for the terrain (Teutoburger Wald and Bannockburn.) One was definitely the result of tactical overconfidence (Antioch II); and the other three could be characterized in those terms.

XIV.D. Conclusions

As we have stated, these 10 conclusive battles are not a large number when compared to the total number of recorded battles. Even if we have found only a few conclusive battles, it appears moderately safe to postulate that conclusive battles are rare. These battles take on mythic and cultural significance far beyond that expected by their frequency of occurrence.

Furthermore, these conclusive battles seem to primarily occur between conquest goal (or rebellion suppressing) forces and conquest avoiding (or rebelling) forces. The battles fall into two broad categories: hold and die battles and cultural disparity battles. In the first category, the losing force is usually a small one fighting a delaying action with the intent of selling themselves dearly. In the second category, the losing force appears to frequently suffer from tactical overconfidence (and inadequate intelligence is frequent evidence of this).The intent of the loser in the first category, the cultural differences (with misunderstanding and prejudice therefrom) between the forces in the second category and the central theme of con quest/anticonquest seem to clearly characterize conclusive battles. This evidence indicates that, historically, the use of conclusion as a victory standard in Lanchester calculations is not well founded.

In terms of lessons in the military art and contemporary politics, however, there is much to learn from conclusive battles. For the recent Desert Storm campaign, the battle for Kuwait was not conclusive but it was decisive. Clearly, the Coalition forces were technically, but not culturally disparate, from the Iraqi forces. It is easy to hypothesize that the joint nature of the Coalition dispelled misunderstanding and prejudice from cultural differences which could have led to overconfidence from the technological disparity of the forces.

Such a view, ignoring other significant factors, would be simplistic and erroneous. Nonetheless, as American interest, and possible military action, focuses more outside of Europe, cultural disparity between forces will be common. The possibility for overconfidence arising from this disparity increases. This must be avoided lest American forces experience a twentieth (or twenty-first) century conclusive battle we neither want nor can afford.

1. Lanchester, Frederick W., **Aircraft in Warfare: The Dawn of the Fourth Arm,** Constable and Company, LTD., London, 1916.

2. Dupuy, R. Ernest, and Trevor N. Dupuy, **The Encyclopedia of Military History from 3500 B.C. to the present**, Harper and Row, New York, 1977.

3. Eggenberger, David, **An Encyclopedia of Battles**, Dover, New York, 1985.

4. Laffin, John, **Brassey's Battles**, Brassey's Defense Publishers, London, 1986.

XV. Of Captains and Colonels, Of Simulations and Sealing Wax

XV.A. Introduction

With this chapter, we reach a crucial juncture in this book. Up to this point, we have been largely concerned with the simplest class of Lanchester problem, the homogeneous aggregation problem. Admittedly, we have taken excursions from time to time, briefly treating with heterogeneous aggregation while considering attacks on fortified lines, and an interlude with stochastic duels, but by and large, our central concern throughout the chapters leading up to this point has been the simplest class of problems for Lanchester attrition theory, two forces contending in combat with no initial differentiation about the spatial, temporal, organizational, physical, and psychological composition of those forces.

If I may take an analogy from Classical Mechanics in Physics, the problem that we have concerned ourselves with thus far is the analog of the interaction of two point masses. While, in physics, we may consider one point mass, as soon as we exert force[a] on that mass, we must introduce another mass (or its field representation). This is the simplest problem one may consider in Classical Physics, just as the two force, homogeneous aggregation problem is the simplest that we may consider in Lanchester Attrition Theory.

In the study of Classical Mechanics, there are several very good reasons for studying this (simple) class of problem. The most basic one is to gain understanding of the basic principles of the mechanics and how these translate in the behaviors (or trajectories) of the masses. The second reason for considering this simple class of problems is that even in more complicated problems involving more masses, the behavior of the masses, much of the time, tends to bear resemblance to the pair-wise behavior. As an example, consider the motion of an extended body (say a projectile in flight.) If that body is rigid (which the projectile will be most of the time, its the exceptions that tend to prove interesting,) then its behavior is described by six quantities - three describing the behavior of the center of mass of the body, and three describing the behavior of the body about the center of mass. When the behavior of the body about the center of mass is well behaved, that is, is stable, then that behavior becomes uninteresting and as a reasonable approximation, we may only consider the behavior of the center of mass of the body. Thus, the problem, under these circumstances, reduces to a point mass problem, the simplest case.

A third reason why we study this simplest class of problem so much is that many of the problems involving multiple masses cannot be solved analytically. These problems that can be solved analytically tend to have special forms or constraints. We must commonly

[a] Note that I have used force with two meanings in this paragraph. The first is the physics term while the second is the Lanchestrian term.

resort to either approximate solutions or simulation. Many of the approximate solution techniques derive from the simple two mass problems, and the simulation results must often be viewed in the two mass framework for understanding. Thus, the simple two mass problem class is the basis for, if not solving, definitely understanding the multi-mass problem.

The same tends to be true in Lanchester Attrition Theory. The simple two force, homogeneous attrition problem is the basis for understanding more complex, multi-force problems. This is not just an excuse for spending so much time and effort in studying these elegant two force solution methods and the behaviors of the solutions. The insights and knowledge gained from this study is the basis for understanding and gaining insight into the multi-force problems.

If you, the student, do not feel confident in your understanding of the theory and methodologies described to this point, do not despair. By its human nature, understanding must grow and frequently does not truly form until we wrestle with problems that challenge us to work with our knowledge and expand it. You should not panic and retire for exhaustive study and contemplation of the material presented thus far, but you should be prepared to return to it selectively on demand, for those purposes. Above all, do not harbor the illusion that we will be leaving this material behind. It is the foundation that we will build on as we proceed and you may likely find occasion to enforce your appreciation of that foundation.

XV.B. The Dog Chases His Tail

To proceed with this transitional chapter, it is necessary that we concern ourselves somewhat deeply with a theoretical consideration of how combat is described. This consideration will seem to bear little resemblance to how the soldier is used to describing, or reading of, combat. I apologize for this, but offer up that once the theory has been described, the soldier will find that it provides a framework that encompasses and accommodates what he is used to, and provides greater scope for the understanding and practice of his skills.

A fundamental concept that we introduce at this time is that of the process. A process is a (possibly complicated) form of activity that results in an accomplishment or product. Obviously, a process must be initiated (begun), and it may conclude (end). If a process is continuous, that is, it continues or repeats indefinitely, then it may not have a conclusion, and it may not be necessary to consider its initiation for the problem under study.

If the activity of the process is the same throughout, then the process is a simple process and cannot be divided further. If the activity of the process is not the same throughout, or different activities are performed during the process, then the process may be subdivided into sub-processes, which may in turn be further subdivided until the condition of sameness is reached. This is, subdivision occurs until the process is fully described in terms

of simple activities.

The composition of a process and its description for any problem is to a certain extent dependent on the problem. For some problems a process may be approximately treated as simple even when it is not. As an approximation then, we will assume that a process may be aggregated, that is, the composition of its subprocesses may be combined in some manner and thereafter ignored, with some (hopefully small) penalty for inaccuracy and some (hopefully large) benefit for simplicity.

This discussion is wonderfully heady and abstruse, but for many of us, the air may now be full of dandruff flakes, if not actual tufts of hair. Let us consider an example, that of the process of breathing. We know that this process is initiated (by birth) and concluded (by death), but most of the time we may safely ignore both of these and consider the process as continuous.

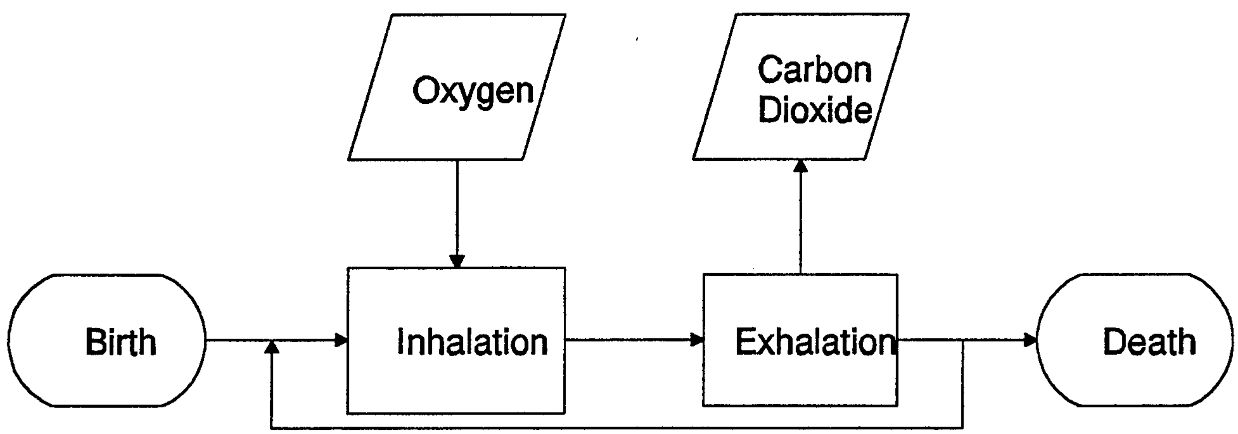


Figure XV.B.1. A Simple Diagram of the Breathing Process

Clearly, the process of breathing is (in the main) repetitive and not simple. It is comprised (under normal conditions) of the repetitive subprocess of inhalation-exhalation which is itself, in turn comprised of two subprocesses, inhalation and exhalation. If we are not concerned with the detailed physiological events which occur during inhalation and exhalation, then we may under certain circumstances, as an approximation, consider inhalation and exhalation as simple processes. In this case, we may effectively describe breathing (under normal circumstances,) as a continuous process comprised of two sequential, simple subprocesses, inhalation and exhalation. The product of the inhalation subprocess is to bring oxygen into the lungs; the product of the exhalation subprocess is to bring carbon dioxide out of the lungs.This process is simply depicted in Figure XV.B.1.

Of course, this is an enormous simplification of the breathing process. It does not address what happens to the oxygen or where the carbon dioxide goes to. It also does not address other variations such as sneezing or coughing, although all of these can be incorporated into the process. For our simple model of the process, we should stick to the

process as we have drawn it for now.

Before proceeding, we need to address some additional variations. First, processes (and subprocesses) may be connected in a variety of ways. Our simple model of breathing has two subprocesses connected in serial. Processes may also be connected in parallel. In addition, products may be positive or negative in the sense that they may be outputs or inputs, which terms may be used synonymously.

In general, processes and their products may be probabilistic (stochastic) or deterministic. It is most convenient to consider all processes and their products as stochastic, and treat the deterministic ones as special cases. By this, we mean that the duration, form or other observables of processes and products have probability distributions associated with them. Further, the connections between processes may also be considered to have probability distributions associated with them. For deterministic processes and products, we can consider these distributions to be delta functions - they have an expectation value, but zero standard deviation.

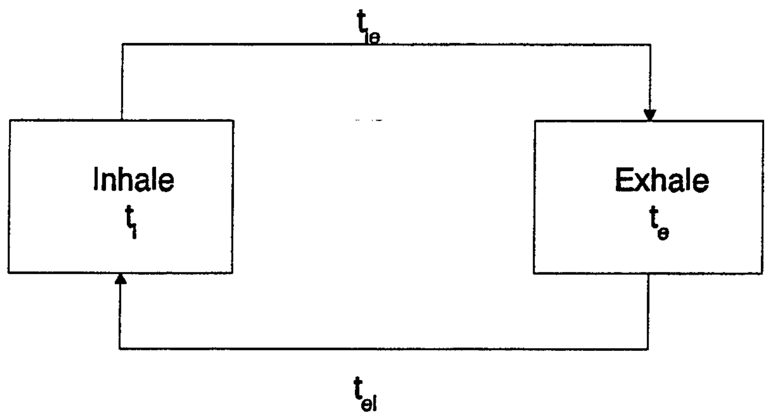


Figure XV.B.2. Continuous Breathing Process with Durations

Quite often in considering combat processes, the most important distributions are those of duration. If we redraw our breathing process diagram as a continuous process (neglecting birth and death,) as shown in Figure XV.B.2 , we see there are four durations involved: the durations of the inhalation (t_i) and the exhalation (t_e) subprocesses, and the durations between the two subprocesses (t_{ie} and t_{ei}). We associate a probability distribution with each of these durations. We may thus associate an expectation value and a standard deviation (and the higher moments) with each of these durations. If we wanted to calculate the expectation value of the duration of one breathing cycle, we would need to consider whether the distributions are independent or not. If they are, then we would only need to sum the expectation values of the four distributions to obtain the cycle's expected duration.

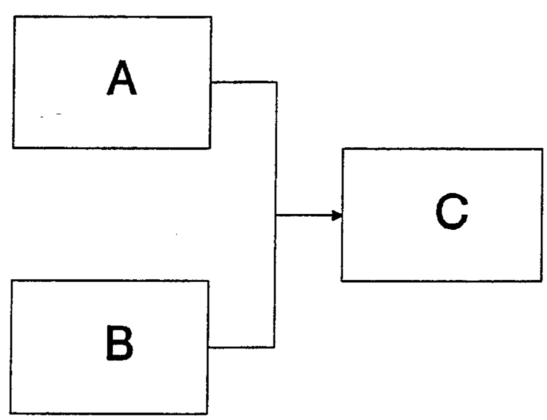


Figure XV.B.3. Contingent Parallel Processes

It also follows for parallel processes that transition or initiation of a subsequent process may be contingent on completion of all or some of the parallel precedent processes. This is shown in Figure XV.B.3.

If both processes A and B must conclude before process C can initiated and the expected durations of these two processes are $\sqsubset t_a \unrhd$ and $\sqsubset t_b \unrhd$, then the expected time before process C can initiate is MAX ($\sqsubset t_a \unrhd$, $\sqsubset t_b \unrhd$).

We must also consider processes that may repeat but which are not continuous. These processes may repeat but they may not be self-repeating. In its simplest form, this may be the result of a decision process. A decision process may have several products, some of which may be the repetition of other processes. We use the convention that the products of a decision process can only be informational in nature. A simple example of this is shown in Figure XV.B.4.

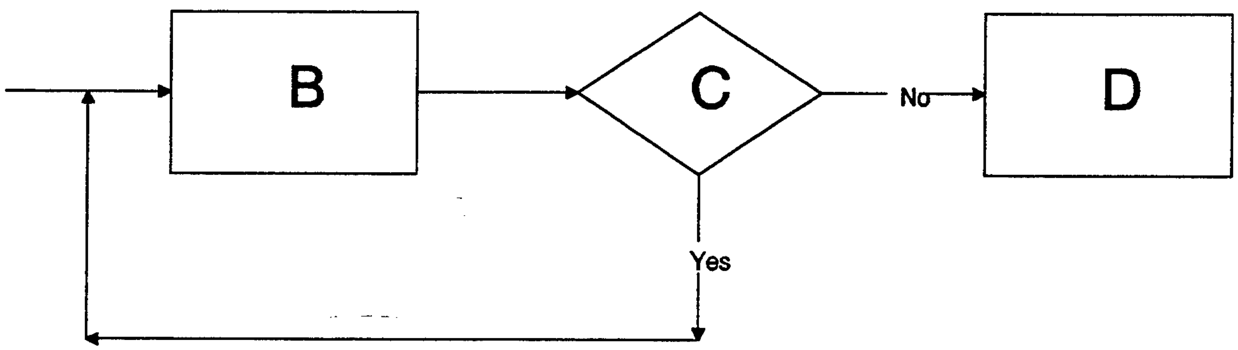


Figure XV.B.4. Repetition through a Decision Process

In this case, process B may repeat as a product of decision process C, or process D may be initiated.

Of particular concern in the description of processes is the nature of the statistics associated with the process. If the process always behaves the same way when it is initiated, then we say that it is a stationary process (even though something happens) because its statistics are stationary. If the process behaves differently, either after the first time it is initiated, or every time it is initiated, we say that it non-stationary because its statistics are non-stationary. Under some circumstances, we may choose to approximate non-stationary processes as stationary ones to simplify the model.

To summarize what we have said about processes:

XV.C. And then lies down

Having spent considerable time talking about the physiology of processes, and considering the breathing process in more detail than any reasonable person would want, you,

the student, are probably asking, "But what does all this have to do with Lanchester Theory?" Ignoring the fact that we chose a process example that every student could identify with (except for those of you who have either just concluded your breathing or reading-this-book processes, and you aren't reading this anyway,) the answer is very simple - combat, battle, war, engagements are processes, and it is within the framework of processes that we may understand, model, and analyze the physics of the battlefield. Indeed, this relationship between the conceptual and descriptive framework of processes and warfare is so tight, so effective, that after a bit of use it becomes difficult to see combat as other than a process of processes. While this is generally true, we must still exercise a certain amount of caution with adopting this as an unqualified world view.

-: A process will normally be initiated and concluded although we may choose to ignore one or both of these if it is continuous or repetitive.

-: Processes have inputs and outputs.

-: Processes are observable; they normally have durations.

-: Decision processes produce only information, although they may pass through products of other processes.

-: Processes may occur in series or in parallel.

-: Processes have probability distribution functions associated with their observables.

-: Stationary processes have stationary statistics.

-: Deterministic processes have delta function distributions.

First of all, we must not loose sight of the fact that any model, and with that, any modeling technique has its limitations. This is true for the process modeling technique. Further, just because we may observe a process and describe its observables, this does not mean that we may quantify the process, either in an absolute sense, or in a relative, approximate sense. Even if we can quantify the process to a degree of accuracy, this does not mean that the model of the process can be made acceptable, either on a technical or a political basis.

As an example of this, consider morale. The morale of a unit may be considered as a process, although admittedly this model is usually of a continuous or frequently re-entrant, non-stationary decision process. There are two fundamental problems with such models however, one of which is technical and the other political. The technical problem with morale models is that they tend to be very difficult to validate (to say nothing of constructing) largely because the observables, both inputs and internal to the process are inherently difficult to quantify and to correlate with outputs. In contemporary terms, the models tend to exhibit chaotic behavior. (And we should expect this, primarily because human beings are involved.) As a result, combat models that incorporate morale models often have surprising or even chaotic results.

The second problem with morale models is political. Within the framework of any organization, it is difficult for the management of that organization to accept any analysis predicting that the organization will perform in a chaotic manner. This is particularly true of military organizations with their inherent focus on mission accomplishment, and their political supervisors who are always asking why they should bear the cost of the military. Accordingly, and as a rule, most military analysis is done without implicit consideration of morale - units are assumed to always perform as ordered.

This does not mean that morale is a process which cannot be considered or even that it cannot be analyzed. But because the quantified morale process model frequently exhibits chaotic results whose consequences in the context of combat analysis are manifestly unacceptable, these analyses are performed without quantified consideration of morale.

At first glance, this deliberate exclusion of the morale process appears to be both technically and ethically unconscionable, but this is not completely the case. From a technical standpoint, the inclusion of the morale process chaotically drives the results. Since we will always be trying to cast analytical results into a framework of stationary statistics for the purpose of understanding, the most common effect of morale models is to increase the variance (i.e., standard deviations) of the calculations.

From a military-political sense, the soldier knows that combat is chaotic, but that soldiers tend to rise above that chaos. Good leadership tends to avert the chaotic results often predicted by the models. Certainly, history, while it records battles where morale broke catastrophically, records them as a small percentage. The rule of history seems to be that widespread and catastrophic unit breakage occurs after the decision, not to cause it. Thus, with morale results as outliers both technically and militarily, analyses without morale consideration can be accepted within the consideration of experimental design, given that morale is considered separately within another methodological framework.

We must emphasize that while there are combat processes that do not lend themselves to an analytical representation in a quantifiable sense, these processes do not occur with such frequency or diverseness to preclude the use of the process modeling technique as a general tool for analyzing combat. As I indicated earlier, the association of process methodology and

combat conceptualization is so natural that after a while "combat process" seems to become one word.

To illustrate this, let us consider the dynamics (and in some cases, statics) of battle. If we consider what a unit goes through in a meeting engagement, then we may describe its dynamics by a set of processes. First, the unit will be moving and this is a process. The movement process, depending on how we model it (and this probably depends on doctrine, type of unit, and time scale of our analysis), may be either continuous or repetitive. We can easily visualize that the movement process may be proceeded by a process of "forming up" and that these two processes are connected by a decision process. The movement process, in turn, may be concluded or altered by a variety of events such as the break down of a vehicle, a rest stop, or even contact with the enemy. In general, we may find it useful to distinguish between movement processes before and after contact with or detection of the enemy.

During the movement, other processes will be occurring in parallel, in particular a scouting or searching process. This process may be carried out by all of the unit or just some of it. Additionally, communications processes, both within the unit, and without, may/will be occurring. Other processes may also be taking place.

Since this is a meeting engagement model, there are three things that may occur when two opposing units come close together in space and time. These are:

-: neither unit detects the other and the movement process continues,

-: both units detect the other,

-: only one unit detects the other.

As we have indicated, if neither unit detects the other, then they will most likely continue in their respective movement processes.

If both units detect each other before combat can occur, then both will probably go through a decision process to decide on the appropriate action. This may include combat, continued movement in a different direction (deferral of combat) or some other action. If both sides decide on combat, or if one unit decides on combat and can close with the other unit, then combat occurs.

If only one unit detects the other, then the possibility for surprise exists. The detecting unit may decide to initiate combat, and do so successfully before it is detected by the other unit. Alternately, the detecting unit may decide to avoid combat, and try to withdraw or otherwise remain undetected. In any case, once detection has occurred, an extra unit communication process may be initiated.

Assuming that the unit decides to engage in combat, then there will probably be a reforming process as the unit changes its formation (say from column to line or from road march to a combat formation). Then, unless it will defend its current position, the unit will

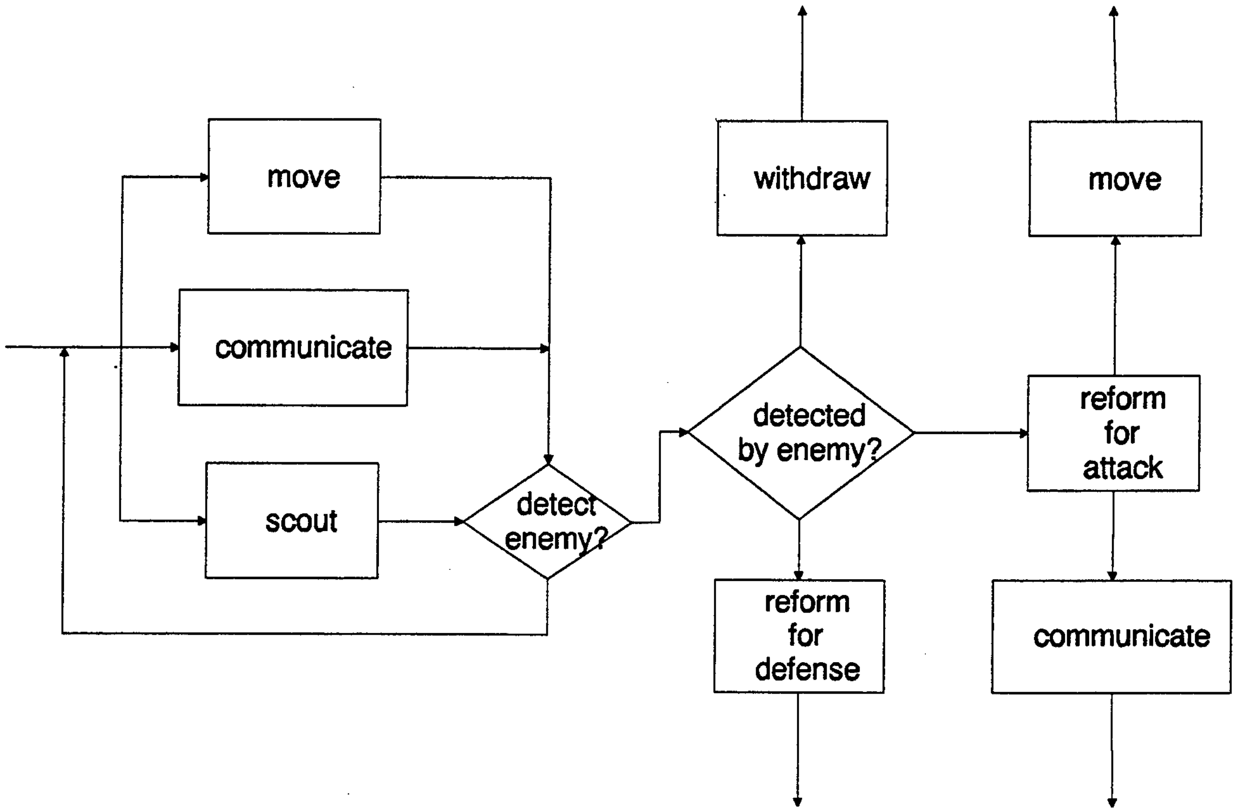


Figure XV.C.1. Meeting Engagement Process Diagram

move towards the enemy unit. This combat movement will probably be a different process than the previous one, and the unit may move as a unit, as sub-units, or even as individuals.[b] Notably, other processes such as acquisition of targets and firing at targets may be occurring serially or in parallel with each other and with movement. Communication may be occurring. This overall process is sketched in Figure XV.C.1.

In general, the dynamics of combat may be described in terms of processes. The description of combat by the description of these processes and their interrelationships is the basis for the analysis of combat in general. It provides the underlying structure for most, if not all, combat models and simulations, and as we shall see, is fundamental to the continued development of Lanchester Attrition Theory.

XV.D. The Master's Voice

Now we come closer to the meat of the discussion, to the subject of aggregation. To address this however, we must first consider some realities of both military and simulation organization. We will consider the latter first.

---

[b] This previous consideration is a fundamental one of de-aggregation.

Our previous discussions have already dealt with the difference between a model and a simulation. Basically, a model is an informational representation of something which is normally, at least in principle, observable while a simulation is a tool, built from one or more models, for the purpose of gaining information about those models and possibly thereby, of reality. A mathematical equation, along with the definition of its terms may be a model, but the instant we take pencil, paper, and calculator and begin to calculate numbers or draw curves based on or using that equation, we are using simulation.

Obviously, the distinction between model and simulation may be a fine one but it is one that we must retain.

One distinction between model and simulation is that a simulation always incorporates at least one model. Actually, the simulation really incorporates at least two models, one being the model environment that the simulation is constructed in. In our example above, this environment was the material pencil, paper, and calculator, the parameters of language, mathematics, etc. In this context, we see that merely writing down a model or speaking it is a simulation.

Why have the class of models at all? Why not just consider everything as a simulation? These are good questions. We want to have the class and concept of model because in one sense a model represents an idealization of information. In another sense, we need it because a model is a product while simulation is a process, albeit we tend to use it as a noun.

Now that we have beaten the metaphysical drum (no pun intended) about simulation and models, we may now move on to discuss some of the anatomy and physiology of simulations. This is the basic reason why we had to through all the discussion about simulations incorporating a model of the environment of the simulation.

Simulations must be expressible in some framework for manipulation. At its simplest, this framework may be thought itself if all we do is think about the model(s). Since the brain/mind is not very good at keeping a large quantity of items in forethought at once, we may be quickly forced to a symbolic representation, say on paper, or to a numeric representation, in a computer. When we do this, we force an environment on the form of the simulation.

In general, we may generally admit of four different kinds of simulation, based on two different descriptions of how the model/simulation behaves. First, the simulation may be *stochastic* or *deterministic*, and deterministic here includes the special case of simulations of stochastic models that only considers the expected values and not the variances - in some sense. This is really a broader class than we might expect because of the different ways that the expected values may be considered. Accordingly, these simulations may sometimes be called *expected value* simulations. Under the appropriate conditions, the Lanchester attrition equations may be viewed as expected value.

The other description of how the simulation works is what is called *time* and *event sequencing*. In time sequencing simulations, time (and space) are the basis of control of the simulation and events occur continuously while time, at least on a computer, is incremental. In event sequenced simulations, the occurrence of events is the basis of control of the simulations; events are discrete and/or sequenced and time and space are continuous. Thus, in a time sequenced simulation, we increment time, pausing at each increment to collect events that have occurred (or concluded,) during the latest increment; in an event sequenced simulation, we order events as they occur, collecting the advance of time as the events occur.

It is possible to describe a simulation as both time and event sequenced if, for example, the incremented passage of some amount of time, without any overt event, is treated as an event. Rigorously, any digital computer based simulation must be both because of the quantified nature of the simulation environment.

If we now make a very large leap and stipulate that, in the main, the model(s) that are being simulated are models of processes (that is, we have a process simulation), then the distinction between time and event sequencing lies in whether we describe the process(es) in terms of their own internal, or of some other external framework.

The choice of framework may be forced on us as a direct consequence of our particular simulation environment (e.g., programming language), or analytical goals, but in general, whether we have a stochastic or determination, time, event or *combined* sequenced simulation, is then not the result of the models that comprise the simulation themselves, but more likely the result of conditions external to the models, including the preferences and prejudices of the simulation developer.

In other words, two simulations of the same processes, one stochastic and event stepped, and the other deterministic and time stepped (e.g., Lanchestrian,) may produce different results (but should not in principle), because the representations of the models in the framework and simulation environment of the simulations are different.

What makes this important for combat modeling is that these two simulations, which simulate the same processes, start from the same process models. This addresses one of the fundamental arguments about Lanchester Attrition Theory, which is that its simplicity and form prevent it from being useful and correlating with other models of combat. If we build our Lanchester based model incorporating these processes (or rather their model representations,) then we may expect some correlation with a higher resolution stochastic (or expected value?,) simulation embodying the same processes. In later chapters, we shall take up the question of correlation by developing the conjugate theory in terms of the processes of combat. For now, we merely illustrate this in Figure XV.D.1, presenting a process comparison between a high resolution stochastic simulation (left,) and a low resolution (aggregated) expected value simulation (right)..

XV.E. "You can't tell them apart without a Score Card".

The terms military and organization are almost inseparable largely due to the high degree of organization that military units have (obviously the use of the term unit will be somewhat confusing here as we must mix an organizational meaning with a Lanchestrian meaning). There is a tendency to see military organizations as blobs (after all, all those soldiers look alike), although there are considerable differences between and within units. For example, an armored or a mechan ized infantry unit may be comprised primarily of tanks or infantry fighting vehicles (and the troops who man them), but may also include other types of weapons systems and troops, and the combat processes, their interrelationships, and specific process forms may thereby be different. Indeed, we may argue that these will be different for two individual troops even when they are in the same combat environments, just because they are individuals.

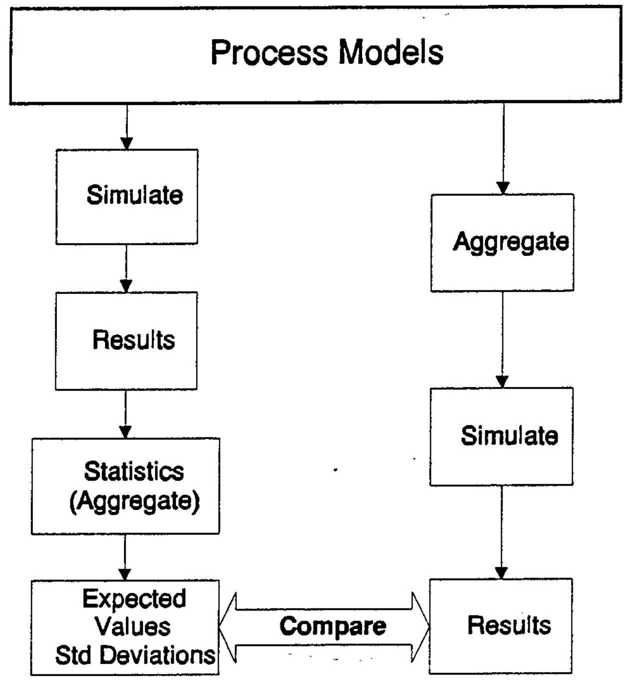


Figure XV.D.1. Comparison of Simulation Methodology

Having noted that military (organizational) units have different structures in terms of their men, equipment, missions, and combat processes, we now proceed to note that these units occupy space and time. Because of the variations due to terrain, the diurnal cycle and weather, these space-time coordinates are not necessarily equal. That is, on rough terrain, at night, or in foul weather, finding enemy units is more difficult than on smooth terrain on a clear day. This also applies to a unit that may be partly on one type of terrain, or in one type of weather, and another unit that is on different terrain or in different weather. Thus, we may consider distinguishing the pieces of this unit on the basis of their trajectories.

Now we may turn to the basic question of aggregation. Fundamentally, aggregation is an ordered process by which certain observables are removed from a model as distinguishing characteristics. For example, in a company of ten tanks, we may consider the company to be comprised of ten units, each a tank, and track the individual behavior of each tank (unit); alternately, we may aggregate the company as a single unit with a strength of ten (tanks,) and track the statistical behavior of a single tank. This tends to simplify the use of the model (i.e., simulation) at the price of some amount of accuracy. Clearly, this is a somewhat arbitrary process whose accuracy (and thereby validity) is assessable only by detailed comparison of simulations incorporating both aggregated and non-aggregated models;

a comparison that is often not made. Contrarily, we may state both idealistically and practically that no simulation has ever been made that does not incorporate at last a model with some degree of aggregation.

For example, we might choose to build a simulation at what we might think of as the lowest logical level - the individual soldier. In this simulation we would simulate the processes and interrelationships all the way down to individual soldiers, but no further. While no major combat simulation has been written with this level of resolution, numerous simulations have been written at the individual weapon system/platform level, usually with dismounted infantry unit aggregated at fire teams or squad level with either discrete or aggregated weapons. An example of this type of simulation is JANUS.

From an historical standpoint, the original rationale for aggregation was the available framework of simulation. Prior to digital computers, stochastic simulation was feasible only for a few degrees of freedom (observables,) and highly tedious (and error prone,) then. Thus, aggregation, usually along organizational lines, was necessary to simulate engagements, battles, and campaigns. This is the genesis of war games.

With the advent of digital computers after World War II, automated repetition became possible; more ambitious stochastic simulations became practicable. With advances in the speed of calculations, the size of memory, and the richness of programming languages, it has become increasingly possible to simulate larger organizations with greater resolution. The trend has been to build division and corps simulations at the individual platform level of resolution. As we have indicated, this approach is embodied in simulations like JANUS, which, during DESERT STORM, was exercised at corps level.

This approach gives us greater confidence in the results of the simulation implicitly, but obviously, there are two difficulties with this, both practical. First of all, the more internal degrees of freedom such a simulation has, that is, the greater the resolution of units and the higher the organizational level - the greater the number of units, the more times the simulation must be exercised for statistical convergence (if not statistical significance!) This has dire consequences when we have several (or many) variations to consider. Second, the mere stressing of the computer resources makes these simulations more dear to exercise because of the cost of the state-of-the-art computer and its capabilities. Thus, even today, there is a continued place for aggregated simulations either as quick analysis tools or as precursors to the larger, less aggregated simulations.

One type of aggregation is on the basis of a military (organizational) unit. This is organizational aggregation and it may occur in one of two ways, depending on how its resources are modeled. Since we are primarily concerned with attrition, we will use weapon examples here although we could make comparable examples in terms of communications or logistics.

Since the unit will have, in general, different weapons systems with different target

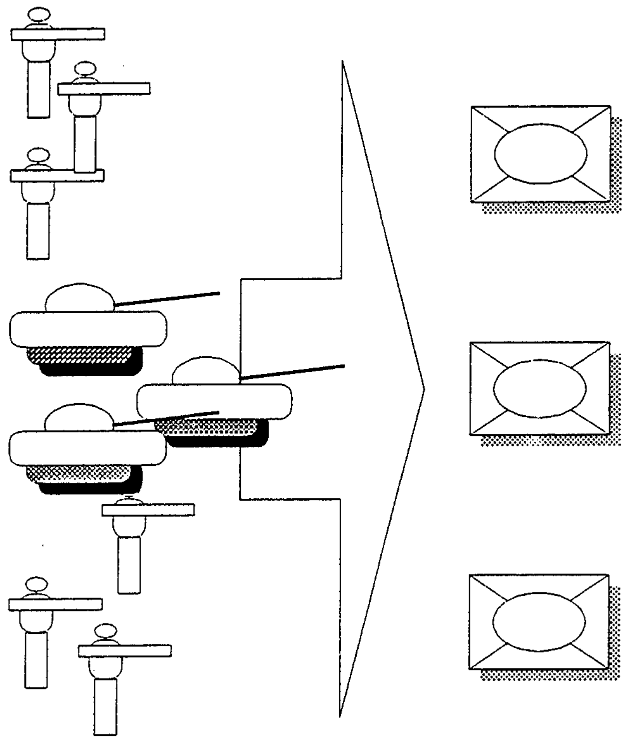


Figure XV.E.1. Aggregation across Weapons Types

engagement characteristics, then if we aggregate all the internal processes at a given unit level, we may either reduce all weapons systems to a common equivalent and adjust either the equivalent force strength (number of Lanchestrian units in the organizational unit), or the relative attrition effectiveness per unit. In general, this type of aggregation has the disadvantage that it fixes the proportions of types of weapons in the unit regardless of losses. This is shown in Figure XV.E.1 where we have taken an idealized infantry-armor task force of three units and aggregated it into three units of uniform weapons.

An alternative type of aggregation is to aggregate all of the weapon systems of a given type. This type of aggregations has the disadvantage that it loses distinctions of losses with respect to organization. This is depicted in Figure XV.E.2 where we have taken our same idealized task force and aggregated into three units, each with uniform weapons. We shall note, at this point without further consideration, that the attrition rates of and against these two types of aggregation are fundamentally similar but completely different in final form and value..

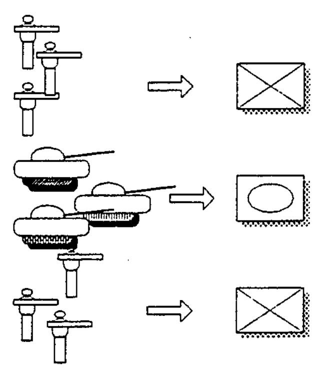


Figure XV.E.2. Aggregation by Weapons Type

A third form of aggregation is to aggregate all of the elements that are in a

particular combat process. This type of aggregation may also lose unit identification. The complexity of this form of aggregation precludes drawing a simple picture at this point.

We may also aggregate on the basis of space and time, especially in a digital computer simulation, since these will have to be discretized in a time or combined sequential simulation. (Obviously we may also event aggregate, which is just a form of process aggregation).

In general, if we aggregate with respect to everything, then we reduce the problem

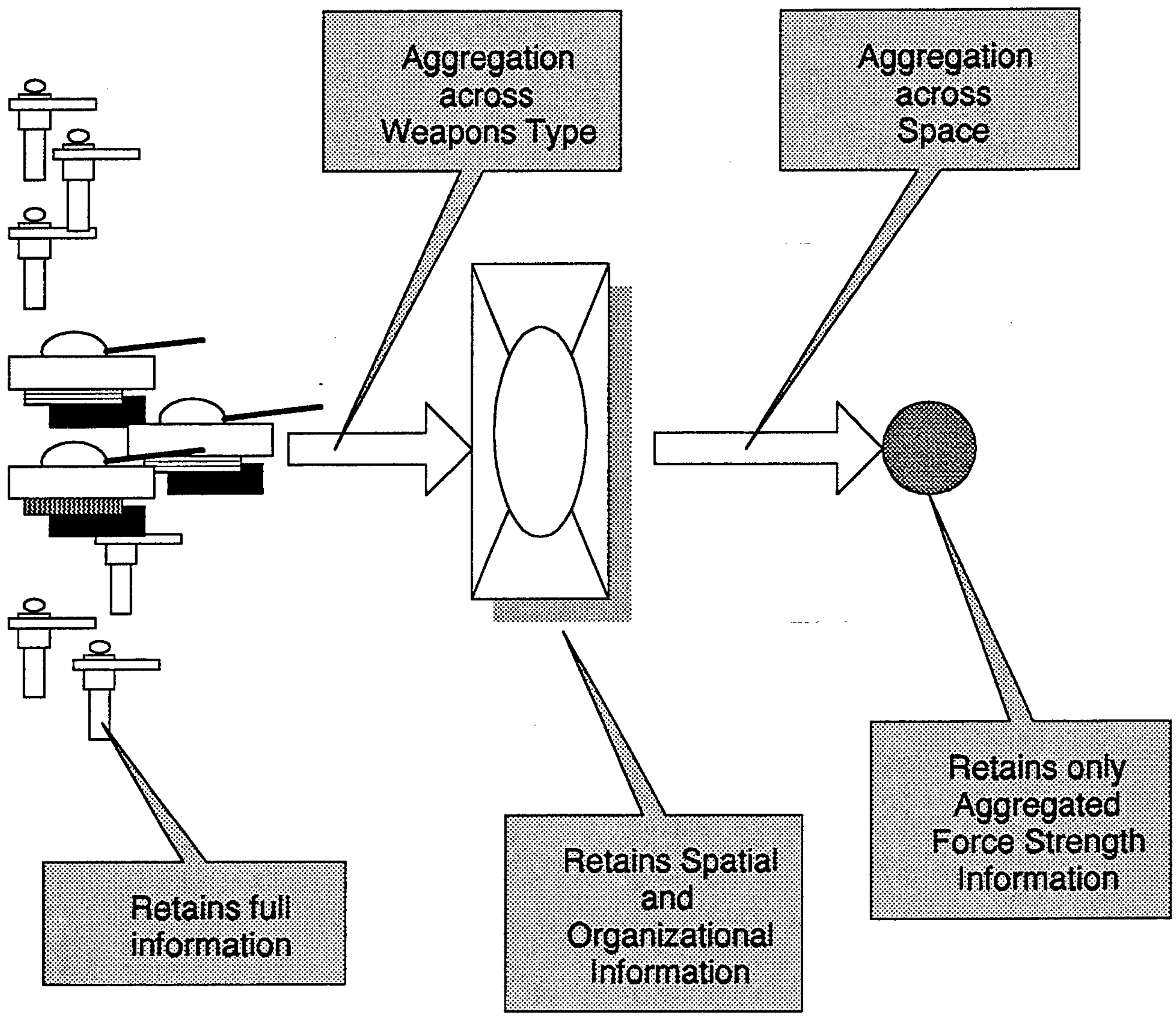


Figure XV.E.3. Aggregation at Total Unit level to Full Homogeneous Aggregation

to one of our familiar homogeneous aggregation where we may have lost all (or at least most) of the distinction among organizations, weapons, etc. This is the case that is sometimes solvable and is supported by the general extent of most historical data. Any other type of aggregation is considered to be heterogenous, although as we have indicated this is now a

very general term. This situation is depicted in Figures XV.E.3, and XV.E.4, which demonstrate two different approaches to further aggregation. In Figure XV.E.3, we depict aggregation of our idealized task force into (first,) a single uniform unit, and then spatially into a single point unit. In Figure XV.E.4, we aggregate our idealized task force, as before, into three weapons uniform spatially dispersed units, and then into a non-uniform point unit. Note that this unit is <u>not</u> homogeneously aggregated.

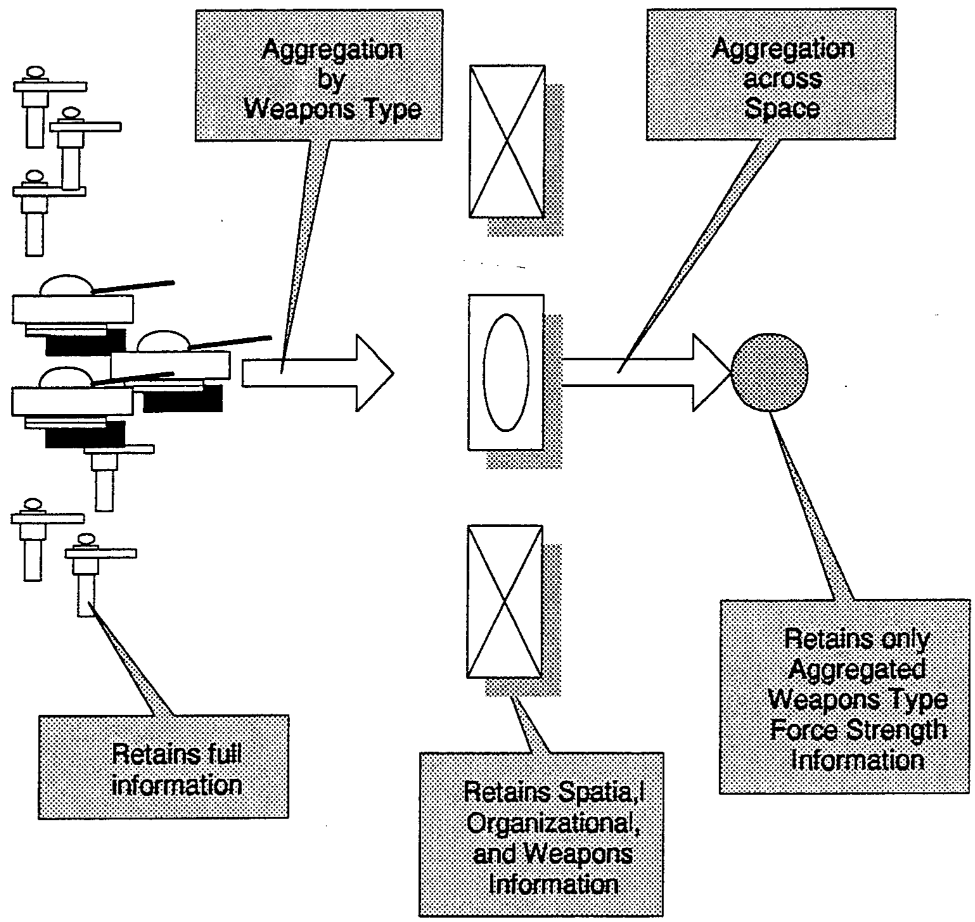


Figure XV.E.4. Aggregation retaining weapons type to spatial homogeneity

The hierarchy of aggregation is a direct relationship between high and low resolutions from counting individual weapons systems through an entire force. As long as we keep account of the combat processes involved, then, we may expect to be able to show some degree of correlation between high and low resolution representations and thereby establish a hierarchy for gaining understanding from the use of these simulations.

XV.F. Aggregation and De-Aggregation

Because of different interests for analysis, it may be desirable to have variable

aggregation or variable resolution models. These models, and simulations from them, are possible, but they are often quite specialized.

For example, let us consider a conceptually simply variable aggregation model/ simulation. We are primarily interested in the simulation of the combat processes so we wish to develop a simulation that aggregates combat elements at the individual weapons platform level in combat, but aggregates elements at an organizational level (say battalion), otherwise. If we consider the primary combat processes for a battalion to be movement, reconstitution and supply, and engagement, then we would treat the first two of these processes at the battalion level, but the engagement process with sub-processes down through the individual platform level. From an aggregation standpoint, this variable aggregation simulation cannot be as highly aggregated as would normally be the case. In particular, the aggregation with respect to systems can only be carried to the point of counting all of the systems of each (distinguishable) type in the battalion so that when engagement starts and the battalion is de-aggregated for combat, the individual system-level aggregation can be reconstructed. In doing this, we may lose detailed sub-battalion force strength information within the reconstitution model approximation, sacrificing consideration of unit integrity and personnel bonds in combat.

Variable resolution aggregation is effected in a similar manner. Say we are primarily interested in one particular class of unit or weapon system. An excellent example of this would be air defense weapons in a division. We might aggregate all of the other weapons or units in the division (or both) at some level, (tank battalions but not the air defense battalions,) but the air defense units/weapons and their targets at some other level. In this case, we would have to pay special attention to the process interrelationships between the air defense units/weapons and other units/weapons. Note also that we probably also have to keep the same level of aggregation for the targets as the firers.

XVI. THE MANYFOLD PATHS

A. INTRODUCTION

With this chapter, we enter a new stage of the book. The general theme of this chapter and the ones immediately following are brief reviews of some of the work that has been performed on extending or embellishing the basic Lanchester attrition formalism. Following this, we will examine several expositions in the context of military operations. Then we shall finally deal with the connection Bonder-Farrell Theory that connects individual level dynamics with attrition mechanics.

At this point, then, we turn a corner. Heretofore, we have dealt with the basics of Lanchester Attrition Theory. Now we push on into the more advanced elements of the theory and its adjuncts. In keeping with our goal to keep the mathematics to approximately the level of calculus, we shall neglect certain topics in detail. Some or all of these topics may be of considerable interest to the individual and we shall, where possible, try to maintain connectivity via citation to permit further study. Admittedly, some of this selection will be somewhat arbitrary, but that is a right we reserve as author. The student should not assume because we slight a topic that it does not have relevance. The field is extensive, and in general, we shall minimize topics only if they require knowledge of some other disciplines beyond the keen of this book.

Having said all of that, we now embark on a review of some other forms of attrition equations. We have, of course, already covered the detail the basic Quadratic and Linear Lanchester Attribution equations, and their general form, the Osipov equations. The Mixed Lanchester Equations, used to describe Guerrilla Warfare have also been covered. All of these equations have a common form, namely

$$\frac{dA}{dt} = -\alpha\, A^{2-n}\, B\,, \qquad \text{(XVI.A-1)}$$

and

$$\frac{dB}{dt} = -\beta\, B^{2-m}\, A\,, \qquad \text{(XVI.A-2)}$$

where n and m are attrition orders. In this chapter, we will now broaden our scope to encompass attrition differential equations of more general form.

B. CONSTANT RATE ATTRITION

The simplest (mathematically) of these "additional" attrition equations is the case of constant attrition. The attrition differential equations have the form:

$$\frac{dA}{dt} = -\alpha, \qquad \frac{dB}{dt} = -\beta , \tag{XVI.B-1}$$

with obvious solutions obtained by direct integration of

$$A(t) = A_0 - \alpha\ t , \qquad B(t) = B_0 - \beta\ t . \tag{XVI.B-2}$$

These equations possess a state solution

$$\alpha \left(B(t) - B_0 \right) = \beta \left(A(t) - A_0 \right) , \tag{XVI.B-3}$$

that is identical to the Lanchester Linear Law.

These equations describe combat that is characterized by a constant rate of loss to both sides such as would be the case on a terrain constrained battlefield (e.g. a bridge or mountain pass) where only a few units on each side may fight. This attrition type has a special parallel to Lanchester attrition with reinforcements when the reinforcements are constrained to mountain a constant number of units in combat.

Since the relevant Quadratic attrition differential equations are

$$\frac{dA}{dt} = -\alpha\ B + a(t) , \tag{XVI.B-4}$$

and

$$\frac{dB}{dt} = -\beta\ A + b(t) , \tag{XVI.B-5}$$

We constrain A(t) and B(t) to be constant (represented by A_c and B_c) until the reinforcements (reserves) are exhausted. This means that the reinforcement rates are simply

$$a(t) = \alpha\ B_c , \tag{XVI.B-6}$$

and

$$b(t) = \beta A_c , \tag{XVI.B-7}$$

subject to the constraints that

$$\int_0^{t_a} a(t') \, dt' = A_r , \tag{XVI.B-8}$$

and

$$\int_0^{t_b} b(t') \, dt' = B_r , \tag{XVI.B-9}$$

where A_r and B_r are the total reinforcements (reserves,) and the total force strengths are just

$$A_T = A_c + A_r , \tag{XVI.B-10}$$

and

$$B_T = B_c + B_r . \tag{XVI.B-11}$$

Since the reinforcement rates are constant, we may easily compute the reinforcement exhaustion times for both sides by integrating equations (XVI.B-8) and (XVI.B-9), and rearranging to yield

$$t_a = \frac{A_r}{\alpha B_c} , \tag{XVI.B-12}$$

$$t_b = \frac{B_r}{\beta A_c} . \tag{XVI.B-13}$$

If we make use of the step function, previously defined in the discussion of the model of attack on fortified lines,

$$\eta(x) = 1 \,,\ x > 0$$
$$= 0 \,,\ x < 0 \,, \qquad \text{(XVI.B-14)}$$

then we may rewrite equations (XVI.B-4) and (XVI.B-5) in explicit form as

$$\frac{dA}{dt} = -\alpha\, B + \alpha\, B_c\, \eta\left(\frac{A_r}{\alpha\, B_c} - t\right), \qquad \text{(XVI.B-15)}$$

and

$$\frac{dB}{dt} = -\beta\, A + \beta\, A_c\, \eta\left(\frac{B_r}{\beta\, A_c} - t\right), \qquad \text{(XVI.B-16)}$$

with the requirement that $t \geq 0$, and initial conditions $A(0) = A_c$ and $B(0) = B_c$.

It is probably not obvious that equations (XVI.B-15) and (XVI.B-16) are equivalent to the constant rate equations (at least for $t \leq \min(t_a, t_b)$). If we define the total force on the battlefield as A_B and B_B, then for $t \leq \min(t_a, t_b)$, these are defined by

$$A_B(t) = A(t) + \int_t^{t_a} a(t')\, dt' \,, \qquad \text{(XVI.B-17)}$$

and

$$B_B(t) = B(t) + \int_t^{t_b} b(t')\, dt' \,. \qquad \text{(XVI.B-18)}$$

If we differentiate these with respect to time, we get

$$\frac{dA_B}{dt} = \frac{dA}{dt} - a(t) \,, \qquad \text{(XVI.B-19)}$$

and

$$\frac{dB_B}{dt} = \frac{dB}{dt} - b(t) \,, \qquad \text{(XVI.B-20)}$$

which reduce to

$$\frac{dA_B}{dt} = -\alpha\ B_c\ , \qquad \text{(XVI.B-21)}$$

and

$$\frac{dB_B}{dt} = -\beta\ A_c\ , \qquad \text{(XVI.B-22)}$$

by virtue of our requirement that

$$A(t) = A_c\ ,\ t < t_a\ , \qquad \text{(XVI.B-23)}$$

and

$$B(t) = B_c\ ,\ t < t_b\ . \qquad \text{(XVI.B-24)}$$

These two differential equations are constant rate since A_c and B_c are constrained. Obviously, once $t \geq \max(t_a, t_b)$, Quadratic attrition again occurs, and while $\min(t_a, t_b) \leq t \leq \max(t_a, t_b)$, a mixed attrition occurs.

C. <u>Exponential Attrition</u>

The next type of attrition that we consider is exponential. The attrition differential equations have the form

$$\frac{dA}{dt} = -\phi\ A\ , \qquad \text{(XVI.C-1)}$$

and

$$\frac{dB}{dt} = -\psi\ B\ . \qquad \text{(XVI.C-2)}$$

These equations have a state solution.

$$\psi \ln\left(\frac{A}{A_0}\right) = \phi \ln\left(\frac{B}{B_0}\right), \qquad \text{(XVI.C-3)}$$

or equivalently

$$\left(\frac{A}{A_0}\right)^{\psi} = \left(\frac{B}{B_0}\right)^{\phi}, \qquad \text{(XVI.C-4)}$$

and solutions

$$A(t) = A_0\, e^{-\phi t}, \qquad \text{(XVI.C-5)}$$

and

$$B(t) = B_0\, e^{-\psi t}. \qquad \text{(XVI.C-6)}$$

On first inspection, one may wonder what kind of combat occurs where the rate of friendly losses is proportional to friendly strength. The most common answer is disease! These equations are often used to describe non-combat losses, although as we have seen in the preceding chapter, a set of mixed attrition equations may arise where one side's loses are proportional to the other side's strength while the other side's loses are proportional to its own strength. These cases arise when the first force is much larger than the second so that only a part of the larger force (a part proportional to the smaller force,) can be brought to bear.

Because of the form of the state solution, exponential attrition is often called the logarithmic law.

D. Helmbold's Modification

Robert Helmbold[1] has pursued this idea further, suggesting that the attrition route be modified to reflect these inequities in force strength ratio. To achieve this, he proposed a pair of attrition differential equations of the form

$$\frac{dA}{dt} = -\alpha\, B\, h\left(\frac{A}{B}\right), \qquad \text{(XVI.D-1)}$$

and

$$\frac{dB}{dt} = -\beta\, A\, g\left(\frac{B}{A}\right), \qquad \text{(XVI.D-2)}$$

where g and h are functions representing these inequities. Helmbold advanced that these correction functions should have unit value for unit argument (i.e. equal sized forces,)

$$h(1) = g(1) = 1 \,. \tag{XVI.D-3}$$

Notice that the two functions have inverse arguments. There are two functions since the corrections to each force may be different because of the training and composition of the force.

Helmbold develops an example where h and g are a simple powers of the argument. (This is mathematically the simplest function that will satisfy the requirements on h and g.) Simply put

$$h(x) = g(x) = x^{c} \,. \tag{XVI.D-4}$$

Thus, we get attrition differential equations

$$\frac{dA}{dt} = -\alpha\, B \left(\frac{A}{B} \right)^{c} , \tag{XVI.D-5}$$

and

$$\frac{dB}{dt} = -\beta\, A \left(\frac{B}{A} \right)^{c} . \tag{XVI.D-6}$$

We may rewrite these as

$$A^{-c} \frac{dA}{dt} = -\alpha\, B^{1-c} , \tag{XVI.D-7}$$

and

$$B^{-c} \frac{dB}{dt} = -\beta\, A^{1-c} . \tag{XVI.D-8}$$

Since

$$A^{-c} \frac{dA}{dt} = \frac{1}{1-c} \frac{dA^{1-c}}{dt} , \tag{XVI.D-9}$$

we may define new variables

$$\begin{aligned} A_c &\equiv A^{1-c} , \\ B_c &\equiv B^{1-c} , \end{aligned} \tag{XVI.D-10}$$

that have attrition differential equations

$$\frac{dA_c}{dt} = -\alpha\ (1-c)\ B_c\ , \tag{XVI.D-11}$$

and

$$\frac{dB_c}{dt} = -\beta\ (1-c)\ A_c\ .$$

These are simply Quadratic attrition differential equations with solutions

$$\begin{aligned} A_c(t) &= A_c(0)\ \cosh\big(\gamma\ (1-c)\ t\big) - \delta\ B_c(0)\ \sinh\big(\gamma\ (1-c)\ t\big)\ , \\ B_c(t) &= B_c(0)\ \cosh\big(\gamma\ (1-c)\ t\big) - \frac{A_c(0)}{\delta}\ \sinh\big(\gamma\ (1-c)\ t\big)\ . \end{aligned} \tag{XVI.D-13}$$

These in turn give us the solutions for the force strengths

$$\begin{aligned} A(t) &= \left[\ A(0)^{1-c}\ \cosh(\gamma\ (1-c)\ t)\ \ -\ \delta\ B(0)^{1-c}\ \sinh(\gamma\ (1-c)\ t)\ \right]^{\frac{1}{1-c}}, \\ B(t) &= \left[\ B(0)^{1-c}\ \cosh(\gamma\ (1-c)\ t)\ -\ \frac{A(0)^{1-c}}{\delta}\ \sinh(\gamma\ (1-c)\ t)\ \right]^{\frac{1}{1-c}}. \end{aligned} \tag{XVI.D-14}$$

Note that when c = 1, the equations (XVI.D-5) and (XVI.D-6) degenerate into exponential attrition.

It is possible to investigate these Helmbold functions further by considering the case where one force is much larger than the other. If we assume $A \gg B$, then we would expect attrition differential equations of the form

$$\frac{dA}{dt} = -\alpha\ B\ , \tag{XVI.D-15}$$

and

$$\frac{dB}{dt} = -\beta\ A\ , \tag{XVI.D-16}$$

based on the argument presented in the proceeding section. If we now compare equations (XVI.D-15) and (XVI.D-16) with equations (XVI.D-1) and (XVI.D-2), we see that

$$\text{Limit}_{\frac{A}{B} \to \infty} h\left(\frac{A}{B}\right) \to 1 \,, \tag{XVI.D-17}$$

and

$$\text{Limit}_{\frac{B}{A} \to 0} h\left(\frac{B}{A}\right) \to \frac{B}{A}. \tag{XVI.D-18}$$

Since Helmbold postulated that

$$h(1) = 1 \,, \tag{XVI.D-19}$$

we see that

$$h(x) = 1 \,,\; x \geq 1 \,, \tag{XVI.D-20}$$

and

$$h(x) \to x \,,\; x \to 0 \,. \tag{XVI.D-21}$$

This is not sufficient information to define h(x) explicitly, but some function of the form

$$h(x) \sim 1 - e^{-\kappa x} \,, \tag{XVI.D-22}$$

has the right behavior for κ sufficiently large. Note that this is the form we found in the Lanchester-Poisson equations.

E. Swiss Army Knife Attrition

In this section, which is almost a preamble to heterogeneous attrition, we take up the mix and match or swiss army knife to attrition. The basic idea here is that we have identified and examined a number of attrition mechanisms and that more than one of these mechanisms may apply.

Let us briefly review the forms of attrition rate that we have examined thus far. For this purpose, we will assume that a friendly force, represented by force strength x(t) is opposed by an enemy force represented by force strength y(t). The attrition rate forms are summarized in table (XVI.D.1).

Table XVI.D.1. Summary of Attrition Rate Forms.

Attrition Rate Form	Attrition Type
Z(t)	Independent, normally used to represent reinforcements or reserves.
α X(t)	Exponential, normally represents looses due to disease or desertion.
α Y(t)	Quadratic, normally represents direct fire losses in the slow kill light, or indirect fire when occupied area varies.
α X(t)Y(t)	Linear, normally represents indirect fire when occupied area remains constant, or direct fire in the slow acquisition limit.

It is a simple matter to combine these attrition rate forms. For example, a combination of all of these could give attrition differential equations of the form.

$$\frac{dA}{dt} = -\alpha\,\phi\,B - \alpha'\,(1-\phi)\,A\,B - \eta\,A + a(t)\ , \qquad \text{(XVI.E-1)}$$

and

$$\frac{dB}{dt} = -\beta\,\psi\,A - \beta'\,(1-\psi)\,B\,A - \xi\,B + b(t)\ , \qquad \text{(XVI.E-2)}$$

where: A, B are the force strengths,
α, β are the direct fire attrition rate coefficients for the two sides,
α', β' are the indirect fire attrition rate coefficients,
η, ξ are the disease attrition rate coefficients,
a(t), b(t) are the reinforcement rates, and
ϕ, ψ are the fractions of the B, A force that engage in direct fire.

In effect, (e.g.) ψA of the red force is engaged in direct fire, while (1 - ψ) A is engaged in indirect fire. As we shall see in the next two chapters, it is eminently possible for ψ and ϕ to be time and/or range dependent functions.

In general, equations (XVI.E-1) and (XVI.E-2) do not possess either a state solution or explicit time solutions. Special cases do exist, but we shall not treat them in detail here.

If a(t) and b(t) are smooth functions, then equations (XVI.E-1) and (XVI.E-2), like most attrition differential equations, are relatively well conditioned, and we may use numerical approximation to solve them. (See How to Build a Spreadsheet Simulation in the Appendices.) In particular, we may use the approximation

$$\int_t^{t+\Delta t} \frac{dA(t')}{dt}\, dt' \approx \frac{dA(t)}{dt}\, \Delta t\,, \qquad \text{(XVI.E-3)}$$

so that equation (XVI.E-1) has the approximate numerical solution form

$$A(t+\Delta t) \approx A(t) - [\, \alpha\, \phi\, B(t) + \alpha'\,(1-\phi)\, A(t)\, B(t) + \eta\, A(t) - a(t)\,]\, \Delta t\,, \qquad \text{(XVI.E-4)}$$

and similarly for equation (XVI.E-2). These equations are readily amenable to spreadsheet simulation.

If a(t) and/or b(t) are not sufficiently smooth, then we are generally forced to take smaller time steps Δt to obtain a reasonable approximation in terms of accuracy. This usually requires us to write a code simulation (using an appropriate language such as BASIC, FORTRAN, PASCAL, or C,) with very small time steps Δt. It may be advantageous to use a more accurate, multi-step integration approximation in a bootstrap fashion. That is, we compute $A(\Delta t)$ (and $B(\Delta t)$) with a one point approximation using A(0) and B(0). We then compute $A(2\Delta t)$ (and $B(2\Delta t)$) using $A(\Delta t)$ and $B(\Delta t)$, and A(0) and B(0) with a two point approximation. We repeat this process of using increasing numbers of points of integration until we reached the number of steps of the objective approximation. If we want to use an n step integration approximation, then we must compute $A((n-1)\Delta t)$ and $B((n-1)\Delta t)$ in this bootstrap manner. Thereafter, we use the n step integration approximation exclusively. Thus, we would compute $A(n\Delta t)$ using the n points $A(0),..A((n-1)\Delta t)$, $B(0),..B((n-1)\Delta t)$.

Because Δt is small, we probably do not want to use all of the computed points $A(i\Delta t)$, $B(i\Delta t)$ in examining and analyzing the results. Therefore, we would only retain selected points, say every $m\Delta t$ for our analysis. This is conveniently done by writing code to store these points in a data file of a format (e.g., comma or comma-quote delimited ASCII,) that facilitates import into a spreadsheet or other geographical program.

If a(t) and b(t) are exactly integrable, even if not smooth, then we may modify the one step integration technologies to avoid this. Let

$$\Xi(t) = \int a(t')\, dt' , \tag{XVI.E-5}$$

and

$$\Omega(t) = \int b(t')\, dt' . \tag{XVI.E-6}$$

Then we may write an approximate solution of equation (XVI.E-1) as

$$\begin{aligned} A(t+\Delta t) \approx A(t) & \\ & - \left[\alpha\, \phi\, B(t) + \alpha' \,(1-\phi)\, A(t)\, B(t) + \eta\, A(t) \right] \Delta t \\ & + \Xi(t+\Delta t) - \Xi(t) , \end{aligned} \tag{XVI.E-7}$$

and similarly for equation (XVI.E-2).

There is an alternative if we eliminate the linear terms from equations (XVI.E-1) and (XVI.E-2). They reduce to

$$\frac{dA}{dt} = -\alpha\, B - \eta\, A + a(t) , \tag{XVI.E-8}$$

and

$$\frac{dB}{dt} = -\beta\, A - \xi\, B + b(t) , \tag{XVI.E-9}$$

which include quadratic and exponential attrition terms, and replacement terms. These attrition differential equations may be thought of describing combat between two forces that are sufficiently numerous (dense) that direct fire combat is killing (rather than acquisition) constrained, are subject to disease losses, and receive reinforcements. These attrition differential equations do not possess a state solution, except in the special case where we set a(t) and b(t) to be zero, and $\eta = h\beta$ and $\alpha = h\xi$, where h is an arbitrary integrable function. In this case, these equations have a linear state solution. This special case does not seem very realistic, however, because it only occurs if one force has a high disease rate coefficient (compare to attrition,) while the other has a low disease rate coefficient. A more useful approach is to rewrite these attrition differential equations as

$$\left(\frac{d}{dt} + \eta \right) A = -\alpha\, B , \tag{XVI.E-10}$$

and

$$\left(\frac{d}{dt} + \xi\right) B = -\beta A , \qquad \text{(XVI.E-11)}$$

recalling that we have set a(t) and b(t) to be zero. We may "differentiate" each of these equations and use the other to form a second order attrition differential equation

$$\left(\frac{d}{dt} + \xi\right)\left(\frac{d}{dt} + \eta\right) A = \gamma^2 A , \qquad \text{(XVI.E-12)}$$

where we have replaced $\alpha\beta$ with γ^2. If we now assume a general solution of the form

$$A \sim e^{rt} , \qquad \text{(XVI.E-13)}$$

and substitute this equation into equation (XVI.E-12). Then we may obtain a quadratic equation of the form

$$(r + \xi)(r + \eta) - \gamma^2 = 0 , \qquad \text{(XVI.E-14)}$$

which has roots

$$r = \frac{-(\xi + \eta) \pm \sqrt{(\xi - \eta)^2 + 4\gamma^2}}{2} . \qquad \text{(XVI.E-15)}$$

From these, it is a simple matter of a single differentiation (application of the boundary conditions,) and some algebra to yield solutions of the form

$$A(t) = A_0 \cosh(\rho t)\, e^{-(\xi+\eta)t} - \frac{\alpha B_0 + \xi A_0}{\rho} \sinh(\rho t)\, e^{-(\xi+\eta)t} , \qquad \text{(XVI.E-16)}$$

and

$$B(t) = B_0 \cosh(\rho t)\, e^{-(\xi+\eta)t} - \frac{\beta A_0 + \eta B_0}{\rho} \sinh(\rho t)\, e^{-(\xi+\eta)t} , \qquad \text{(XVI.E-17)}$$

where:

$$\rho = \frac{\sqrt{(\xi - \eta)^2 + 4\gamma^2}}{2} . \qquad \text{(XVI.E-18)}$$

Thus equations (XVI.E-10) and (XVI.E-11) have explicit analytical solutions even if they do not have a state solution because of the exponential decay terms solution. (Remember, we told you this would happen.) These solutions may be used to form solutions of equations (XVI.E-8) and (XVI.E-9) in the same manner that we used the solutions to the basic Quadratic Lanchester attrition differential equations (without reinforcement,) to form the solutions with reinforcements. Because the method is the same, and we have previously discussed it in detail, we shall leave this application of it as an exercise for the student.

F. Peterson's Logarithmic Equations

Weiss[2] reports that Richard H. Peterson[3] has observed that attrition differential equations of the form

$$\frac{dA}{dt} = -\alpha \ln(B) A , \qquad \text{(XVI.F-1)}$$

and

$$\frac{dB}{dt} = -\beta \ln(A) B , \qquad \text{(XVI.F-2)}$$

show agreement with actual data on tank battles. These equations show losses that increase with force committed. Several authors have suggested that a force's vulnerability to loss is proportional to the force's size, but its ability to attrit increases functionally slower. This is exactly the situation we have addressed earlier in discussions following our consideration of Osipov. To briefly review this, consider a Napoleonic unit. Assuming area per element is constant, then the area covered by the unit increases linearly with number of elements. However, only the elements on the periphery can easily use their weapons. This number of elements that may fire is thus proportional to the square root of the total number of elements in the unit. Thus Peterson's equations are another attempt to address this same problem.

We may rewrite equations (XVI.F-1) and (XVI.F-2) as

$$\frac{d\ln(A)}{dt} = -\alpha \ln(B) , \qquad \text{(XVI.F-3)}$$

and

$$\frac{d\ln(B)}{dt} = -\beta \ln(A) , \qquad \text{(XVI.F-4)}$$

which we immediately recognize as having a "Quadratic" law type of form. From this, we may immediately write down solutions,

$$\ln\big(A(t)\big) = \ln\big(A_0\big) \cosh(\gamma t) - \delta \ln\big(B_0\big) \sinh(\gamma t) , \qquad \text{(XVI.F-5)}$$

and

$$\ln\big(B(t)\big) = \ln\big(B_0\big) \cosh(\gamma t) - \frac{\ln\big(A_0\big)}{\delta} \sinh(\gamma t) . \qquad \text{(XVI.F-6)}$$

These may be rewritten by taking the antilogarithm as

$$A(t) = \frac{A_0^{\cosh(\gamma t)}}{B_0^{\delta \sinh(\gamma t)}} , \qquad \text{(XVI.F-7)}$$

and

$$B(t) = \frac{B_0^{\cosh(\gamma t)}}{A_0^{\frac{\sinh(\gamma t)}{\delta}}} .$$

The Peterson equations suffer from one difficulty - they do not possess an Ironman solution since the logarithm of one is zero. Thus, we do not have any means for estimating, or even calculating, on an analytical basis, the attrition rate coefficients. Until some means is developed for calculating attrition rate coefficients for the Peterson attrition equations, this deficit will probably fatally compromise the application of these equations.

References

1. Helmbold, Robert L., "A Modification of Lanchester's Equations", Operations Research 13, pp. 857-859, 1965.

2. Weiss, Herbert K., "Combat Models and Historical Data: The U. S. Civil War", Operations Research 14(5), pp. 759-790, 1966.

3. Peterson, Richard H., "On the 'Logarithmic Law' of Attrition and its Application to Tank Combat", Operations Research 15(3), pp. 557-558, 1967.

XVII. Time Dependent Attrition

A. Introduction

The student will probably be looking at the title of this chapter somewhat wonderingly. After all, the whole book (thus far,) has been about attrition; we have been solving differential equations with time as the independent variable; and that continually makes them time dependent, doesn't it?

The answer, of course, is that all of the attrition rates that we have considered thus far have been time dependent only in the force strengths. The attrition rate coefficients have been constants; not functions of anything. Previously, we made some to-do about attrition rate coefficients really being functions and that constant values were special cases. Well, in this chapter (and the next) we begin to wrestle with the mathematics associated with attrition rate functions.

Many workers have contributed to this area, but the capstone worker has been Taylor,[1,2,3,4,5,6] that most prolific of workers on Lanchestrian Theory. His works on the subject of attrition rate functions of time amply review much of the applicable literature as well as contributing the acme in the area. Sadly, we shall be unable to deal with much of the detail of this work because its mathematical complexity is beyond the scope of this text. The student who is more mathematically inclined should be able to digest his articles readily and may wish to do so as an advanced exercise.

B. Quadratic Attrition

Most of what we shall deal with in this chapter will be akin to Quadratic Lanchestrian attrition although it is a strictly misnomer to call it Quadratic (or Linear) since in most cases, there is no state solution. That is, because the attrition rate functions are functions of time, they cannot in general be removed to form a state solution. We shall tend however to follow a somewhat sloppy convention to describe the attrition differential equations in the same manner that we have thus far when the attrition rate functions are constants.

B.1. The Basic Equations

The basic Quadratic-like attrition differential equations have the same general appearance that they had before,

$$\frac{d}{dt} A(t) = - \alpha(t)\, B(t) \,, \qquad \text{(XVII.B-1)}$$

and

$$\frac{d}{dt} B(t) = -\beta(t)\, A(t)\ , \qquad \text{(XVII.B-2)}$$

except that now, the attrition rate functions are now just that.

We shall place some mathematical restrictions on the attrition rate functions. In particular, we shall require them to be non-negative,

$$\alpha(t)\ ,\ \beta(t) \geq 0\ ,\ \forall t \geq 0\ , \qquad \text{(XVII.B-3)}$$

otherwise they would generate rather than destroy force strength. Further, we require these attrition rate functions to be both differentiable and integrable, and their integral be strictly increasing,

$$\int_0^{t+\varepsilon} \alpha(t')\, dt' > \int_0^{t} \alpha(t')\, dt'\ ,\ \forall t > 0\ . \qquad \text{(XVII.B-4)}$$

It is further desirable that the attrition rate temporal progress, somewhat similar to dimensionless proper time in relativity, defined by

$$\tau(t) \equiv \int_0^{t} \sqrt{\alpha(t')\, \beta(t')}\, dt'\ , \qquad \text{(XVII.B-5)}$$

exist and be strictly increasing.

B.2 A Special Case

If the attrition rate functions have the same functional dependence to within a constant, that is,

$$\alpha(t) = \alpha\ h(t)\ , \qquad \text{(XVII.B-6)}$$

and

$$\beta(t) = \beta\ h(t)\ , \qquad \text{(XVII.B-7)}$$

there the attrition differential equations possess a state solution

$$\frac{dA}{dt}\frac{dt}{dB} = \frac{dA}{dB} = \frac{\alpha(t)\ B(t)}{\beta(t)\ A(t)} = \frac{\alpha}{\beta}\frac{B}{A}, \tag{XVII.B-8}$$

just as if they did when the attrition rate functions were constants. In this case, the attrition rate progress is just

$$\tau(t) = \gamma \int_0^t h(t')\ dt', \tag{XVII.B-9}$$

and the attrition differential equations posses solutions

$$A(t) = A_0 \cosh(\tau(t)) - \delta\ B_0 \sinh(\tau(t)), \tag{XVII.B-10}$$

and

$$B(t) = B_0 \cosh(\tau(t)) - \frac{A_0}{\delta} \sinh(\tau(t)), \tag{XVII.B-11}$$

that look like the solutions we got when the attrition rate functions were constants. Since

$$\frac{d\tau}{dt} = \gamma\ h(t), \tag{XVII.B-12}$$

we immediately see by differentiation

$$\begin{aligned} \frac{dA}{dt} &= \frac{d\tau}{dt}\frac{dA}{d\tau} \\ &= \gamma\ h(t) \left[A_0 \sinh(\tau) - \delta\ B_0 \cosh(\tau) \right] \\ &= -\alpha\ h(t)\ B_0 \cosh(\tau) + \gamma\ h(t)\ A_0 \sinh(\tau) \\ &= -\alpha(t) \left[B_0 \cosh(\tau) - \frac{A_0}{\delta} \sinh(\tau) \right] \\ &= -\alpha(t)\ B(t), \end{aligned} \tag{XVII.B-13}$$

which proves our point.

B.3 A More General Case

In general, $\alpha(t)$ and $\beta(t)$ do not have the same functional dependence,

$$A(t) = \sum_{i=0}^{\infty} A_i \frac{t^i}{i!}, \qquad \text{(XVII.B-20)}$$

and

$$B(t) = \sum_{i=0}^{\infty} B_i \frac{t^i}{i!}, \qquad \text{(XVII.B-21)}$$

Of course,

$$A_i = \frac{d^i A(t)}{dt^i}\bigg|_{t=0}, \qquad \text{(XVII.B-22)}$$

and similarly for the $\mathbf{B_i}$, but we do not make use of this. In complement, we write the attrition rate functions as series

$$\begin{aligned} \alpha(t) &= \sum_{i=0}^{\infty} \frac{d^i\alpha(t)}{dt^i}\bigg|_{t=0} \frac{t^i}{i!} \\ &= \sum_{i=0}^{\infty} \alpha_i \frac{t^i}{i!}, \end{aligned} \qquad \text{(XVII.B-23)}$$

and similarly for $\beta(t)$.

Next, we differentiate equations (XVII.B-20) and (XVII.B-21),

$$\begin{aligned} \frac{dA}{dt} &= \sum_{i=0}^{\infty} \frac{A_i\, t^{i-1}}{(i-1)!} \\ &= \sum_{i=0}^{\infty} \frac{A_{i+1}\, t^i}{i!}, \end{aligned} \qquad \text{(XVII.B-24)}$$

(and similarly for B(t). Then we substitute these equations into equations (XVII.B-1) and (XVII.B-2)to yield (for XVII.B-1)

$$\sum_{i=0}^{\infty} \frac{A_{i+1}\, t^i}{i!} = \sum_{j,k=0}^{\infty} \frac{\alpha_j\, t^j}{j!} \frac{B_k t^k}{k!}, \qquad \text{(XVII.B-25)}$$

and shift the indices on the right-hand side of the equation

$$B(t+\Delta t) \approx B(t) - \beta(t)\, A(t)\, \Delta t \,. \tag{XVII.B-32}$$

These equations are easily put into a spreadsheet. We can make use of more elaborate schemes, as we have previously indicated, but care must be taken to ensure adequate stability.

C. Time Average Approximation

Another approximation that we may use is to replace the time dependent attrition rate functions with their average over an internal. If we define the average of the attrition rate function over the interval (t, t + Δt) as

$$\langle \underline{\alpha(t)} \rangle_{\Delta t} = \frac{1}{\Delta t} \int_t^{t+\Delta t} \alpha(t')\, dt' \,, \tag{XVII.C-1}$$

then there is some error ε,

$$\varepsilon = \left| \alpha(t) - \langle \underline{\alpha(t)} \rangle_{\Delta t} \right| , \tag{XVII.C-2}$$

which takes some maximum value on the interval. If $\alpha(t)$ is strictly monotonically increasing, then the maximum error occurs at one of the two endpoints.

If we make this error sufficiently small for our purposes, then we may approximate the attrition differential equations on the interval as

$$\frac{dA}{dt'} \approx -\langle \underline{\alpha(t)} \rangle_{\Delta t}\, B \,, \; t \le t' \le t + \Delta t \,, \tag{XVII.C-3}$$

and

$$\frac{dB}{dt'} \approx -\langle \underline{\beta(t)} \rangle_{\Delta t}\, A \,, \; t \le t' \le t + \Delta t \,, \tag{XVII.C-4}$$

which, as we know have solutions

$$\mathbf{A}(t + t'') = \mathbf{A}(t) \cosh(\langle \underline{\gamma(t)} \rangle_{\Delta t} t'') - \langle \underline{\delta(t)} \rangle_{\Delta t}\, \mathbf{B}(t) \sinh(\langle \underline{\gamma(t)} \rangle_{\Delta t} t'') \,, \tag{XVII.C-5}$$

and

$$\mathbf{B}(t + t'') = \mathbf{B}(t) \cosh(\langle \underline{\gamma(t)} \rangle_{\Delta t} t'') - \frac{\mathbf{A}(t)}{\langle \underline{\delta(t)} \rangle_{\Delta t}} \sinh(\langle \underline{\gamma(t)} \rangle_{\Delta t} t'') \,, \tag{XVII.C-6}$$

where

$$0 \le t'' \le \Delta t$$

$$\langle \underline{\gamma(t)} \rangle_{\Delta t} = \sqrt{\langle \underline{\alpha(t)} \rangle_{\Delta t} \, \langle \underline{\beta(t)} \rangle_{\Delta t}}$$

$$\langle \underline{\delta(t)} \rangle_{\Delta t} = \sqrt{\frac{\langle \underline{\alpha(t)} \rangle_{\Delta t}}{\langle \underline{\beta(t)} \rangle_{\Delta t}}} \,. \tag{XVII.C-7}$$

These equations allow us to estimated the force strength, anywhere on the interval.

If we now introduce a notation that

$$A(i\Delta t) \equiv A_i$$

$$B(i\Delta t) \equiv B_i$$

$$\underline{\alpha_i} \equiv \frac{1}{\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} \alpha(t') \, dt$$

$$\underline{\beta_i} \equiv \frac{1}{\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} \beta(t') \, dt \,, \tag{XVII.C-8}$$

then we may write approximate solutions

$$A_{i+1} = A_i \cosh(\sqrt{\underline{\alpha_i}\,\underline{\beta_i}} \;\Delta t) - \sqrt{\frac{\underline{\alpha_i}}{\underline{\beta_i}}} \, B_i \sinh(\sqrt{\underline{\alpha_i}\,\underline{\beta_i}} \;\Delta t) \tag{XVII.C-9}$$

and

$$B_{i+1} = B_i \cosh(\sqrt{\underline{\alpha_i}\,\underline{\beta_i}} \;\Delta t) - A_i \sqrt{\frac{\underline{\beta_i}}{\underline{\alpha_i}}} \sinh(\sqrt{\underline{\alpha_i}\,\underline{\beta_i}} \;\Delta t) \tag{XVII.C-10}$$

that allow us to proceed in a bootstrap manner.

Since we are really making use of derived quantities γ and δ, it is more accurate to use the time averaged of these quantities

$$\underline{\gamma_i} \equiv \frac{1}{\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} \sqrt{\alpha(t') \, \beta(t')} \; dt' \,, \tag{XVII.C-11}$$

and

$$\underline{\delta_i} \equiv \frac{1}{\Delta t} \int_{i\Delta t}^{(i+1)\Delta t} \sqrt{\frac{\alpha(t')}{\beta(t')}}\; dt' , \qquad \text{(XVII.C-12)}$$

Although these calculations may be more difficult and therefore prohibitive. If these can be calculated (or computed,) then the solutions are

$$A_{i+1} = A_i \cosh(\underline{\gamma_i}\ \Delta t) - \underline{\delta_i}\, B_i \sinh(\underline{\gamma_i}\ \Delta t) \qquad \text{(XVII.C-13)}$$

and

$$B_{i+1} = B_i \cosh(\underline{\gamma_i}\ \Delta t) - \frac{A_i}{\underline{\delta_i}} \sinh(\underline{\gamma_i}\ \Delta t) . \qquad \text{(XVII.C-14)}$$

D. Linear Attrition

Linear Lanchestrian-like attrition with time dependent attrition rate functions have attrition rate differential equations that may be defined in an analogous manner to equations (XVII.B-1) and (XVII.B-2),

$$\frac{d}{dt} A(t) = -\alpha(t)\, A(t)\, B(t) , \qquad \text{(XVII.D-1)}$$

and

$$\frac{d}{dt} B(t) = -\beta(t)\, B(t)\, A(t) . \qquad \text{(XVII.D-2)}$$

Of course, analytical solutions of these equations are more difficult than for Quadratic attrition, in part because we cannot form the equations comparable to equations (XVII.B-18) and (XVII.B-19). If the ratio of attrition rate functions is a constant, then we may form a solution in an analogous manner to equation (XVII.B-8) since a state solution will exist. Otherwise, the student is again referred to the work of Taylor for analytical solutions.

The method of Frobenius also does not provide much joy, because of the presence of three series on the right hard sides of the equations. We are thus left with two viable alternatives: approximate numerical integration and time averaging of the attrition rate functions. Both of these techniques are applicable and workable, again assuming that $\alpha(t)$ and $\beta(t)$ are well behaved mathematically, and Δt is appropriately chosen. In fact, we would expect time averaging to work better with linear attrition since we do not have to introduce derived quantities

such as γ and δ.

E. A Practical Example

In later chapters, we shall take up the detailed engineering theory of attrition rate functions and see therein how time (and range) dependent attrition rate functions arise. As an example, we shall now consider a simple combat calculation that leads use to time dependent attrition rate functions in a derivative manner.

Consider two forces, red and blue, each of which is comprised of two types of units which conduct direct and indirect fire engagements, respectively. Designate the time dependent force strength components as A_d, A_i, B_d, B_i with subscripts d and i for direct and indirect, respectively. (We have suppressed the time dependence for notational compactness.) Assume the two direct fire components to only engage each other, but the indirect fire components to engage in both support and counter-battery fire. For simplicity, assume that the constants, and homogenous aggregation with engagement type. Then the attrition differential equations are:

$$\frac{dA_d}{dt} = -\alpha\, B_d - \alpha'\,(1-\phi)\, A_d\, B_i\,, \qquad \text{(XVII.E-1)}$$

$$\frac{dB_d}{dt} = -\beta\, A_d - \beta'\,(1-\psi)\, B_d\, A_i\,, \qquad \text{(XVII.E-2)}$$

$$\frac{dA_i}{dt} = -\eta\,\phi\, A_i\, B_i\,, \qquad \text{(XVII.E-3)}$$

and

$$\frac{dB_i}{dt} = -\zeta\,\psi\, B_i\, A_i\,. \qquad \text{(XVII.E-4)}$$

The first two equations describe the attrition of the direct fire units while the last two equations describe the attrition of the indirect fire units (the counter-battery engagement). The first right hand side terms in the first two equations are the direct fire attrition rates while the second terms are the indirect fire attrition rates. The quantities α, α', β, β', η and ζ are the attrition rate coefficients, and the parameters ϕ and ψ allocate fire between the support and counter-battery missions. Obviously, $0 \le \phi,\psi \le 1$. Note that a state solution obviously exists for A_i and B_i (since they are not subject to direct fire,) and that analytical solutions for these two force strength components exist.

If we define

$$\Delta_i \equiv \zeta \ \psi \ A_i \ - \eta \ \phi \ B_i \ , \tag{XVII.E-5}$$

to represent the state solution for the indirect fire forces, then we may write the analytic solutions for these two force strength components as

$$A_i(t) = \frac{A_i(0) \ \Delta_i}{\zeta \ \psi \ A_i(0) \ - \eta \ \phi \ B_i(0) \ e^{-\Delta_i t}} \ , \tag{XVII.E-6}$$

and

$$B_i(t) = \frac{B_i(0) \ \Delta_i \ e^{-\Delta_i t}}{\zeta \ \psi \ A_i(0) \ - \eta \ \phi \ B_i(0) \ e^{-\Delta_i t}} \ , \tag{XVII.E-7}$$

from our previous investigations. Retaining the time dependence notation of these two force components as a shorthand for these two equations, we may now rewrite equations (XVII.E-1) and (XVII.E-2) as

$$\frac{dA_d}{dt} = - \alpha \ B_d \ - \left[\alpha' \ (1 \ - \phi) \ B_i(t)\right] A_d \ , \tag{XVII.E-8}$$

and

$$\frac{dB_d}{dt} = - \beta \ A_d \ - \left[\beta' \ (1 \ - \psi) \ A_i(t)\right] B_d \ . \tag{XVII.E-9}$$

From the standpoint of solving equations (XVII.E-8) and (XVII.E-9), the quantities in brackets (i.e., []) represent time dependent attrition rate functions that derive from the formulation of the problem. Note that these two equations are of mixed Quadratic-Exponential form where the attrition due to indirect fire support takes on exponential form with time dependent attrition rate function.

There are several approaches to solving equations (XVII.E-8) and (XVII.E-9). Because we expect A_i and B_i to be smooth, well-behaved functions of time, both approximate numerical integration and time averaging are attractive. We may also examine solving these equations by a variation of the methods used in the preceding chapter.

If we write trial solutions of the form

$$A_d = a_d\, e^{-\int_0^t \alpha' (1 - \phi)\, B_i(t')dt'}, \qquad \text{(XVII.E-10)}$$

and

$$B_d = b_d\, e^{-\int_0^t \beta' (1 - \psi)\, A_i(t')dt'}, \qquad \text{(XVII.E-11)}$$

then we may eliminate the exponential attrition terms from equation(XVII.E-8) and (XVII.E-9). Before doing this, however, we want to examine the exponential terms in equations (XVII.E-10) and (XVII.E-11).

From equations (XVII.E-3) and (XVII.E-4), we may note that

$$A_i(t) = -\frac{1}{\zeta\psi}\frac{d\ln(B_i)}{dt}, \qquad \text{(XVII.E-12)}$$

and

$$B_i(t) = -\frac{1}{\eta\phi}\frac{d\ln(A_i)}{dt}. \qquad \text{(XVII.E-13)}$$

If we now substitute equation (XVII.E-12) into the exponential factor of equation (XVII.E-11),

$$e^{-\int_0^t \alpha' (1 - \phi)\, A_i(t')\, dt'} = e^{\int_0^t \frac{\alpha' (1 - \phi)}{\zeta\psi}\frac{d\ln(B_i(t'))}{dt'}dt'}, \qquad \text{(XVII.E-14)}$$

and note that the integral is exact, then

$$e^{-\int_0^t \alpha' (1 - \phi)\, A_i(t')\, dt'} = e^{\frac{\alpha' (1 - \phi)}{\zeta\psi}\left[\ln(B_i(t) - \ln(B_i(0)\right]} = \left[\frac{B_i(t)}{B_i(0)}\right]^{\frac{\alpha' (1 - \phi)}{\zeta\psi}}. \qquad \text{(XVII.E-15)}$$

We may now rewrite equations (XVII.E-1) and (XVII.E-2) using equations (XVII.E-10), (XVII.E-11), and (XVII.E-15) (and its equivalent from equation (XVII.E-11)) as

$$\left[\frac{B_i(t)}{B_i(0)}\right]^{\frac{\alpha'(1-\phi)}{\zeta\psi}} \frac{da_d}{dt} = -\alpha \left[\frac{A_i(t)}{A_i(0)}\right]^{\frac{\beta'(1-\psi)}{\eta\phi}} b_d \,, \qquad \text{(XVII.E-16)}$$

and

$$\left[\frac{A_i(t)}{A_i(0)}\right]^{\frac{\beta'(1-\psi)}{\eta\phi}} \frac{db_d}{dt} = -\beta \left[\frac{B_i(t)}{B_i(0)}\right]^{\frac{\alpha'(1-\phi)}{\zeta\psi}} a_d \,. \qquad \text{(XVII.E-17)}$$

For convenience, we introduce the definitions

$$\alpha'' \equiv \frac{\alpha'(1-\phi)}{\zeta\psi} \,, \qquad \text{(XVII.E-18)}$$

and

$$\beta'' \equiv \frac{\beta'(1-\psi)}{\eta\phi} \,, \qquad \text{(XVII.E-19)}$$

and rewrite equations (XVII.E-16) and (XVII.E-17) as

$$\left[\frac{A_i(t)}{A_i(0)}\right]^{-\beta''} \left[\frac{B_i(t)}{B_i(0)}\right]^{\alpha''} \frac{da_d}{dt} = -\alpha\, b_d \,, \qquad \text{(XVII.E-20)}$$

and

$$\left[\frac{B_i(t)}{B_i(0)}\right]^{-\alpha''} \left[\frac{A_i(t)}{A_i(0)}\right]^{\beta''} \frac{db_d}{dt} = -\beta\, a_d \,, \qquad \text{(XVII.E-21)}$$

and differentiate these. (We will carry forward with only one of these for simplicity.) This gives

$$\begin{aligned} \frac{d}{dt}\left[\frac{A_i(t)}{A_i(0)}\right]^{-\beta''} \left[\frac{B_i(t)}{B_i(0)}\right]^{\alpha''} \frac{da_d}{dt} &= -\alpha\, \frac{db_d}{dt} \,, \\ &= \alpha\, \beta \left[\frac{A_i(t)}{A_i(0)}\right]^{-\beta''} \left[\frac{B_i(t)}{B_i(0)}\right]^{\alpha''} a_d \,, \end{aligned} \qquad \text{(XVII.E-22)}$$

which we note simplifies by equations (XVII.E-12) and (XVII.E-13) to

$$\frac{d^2 a_d}{dt^2} + \left[\beta' (1 - \psi)\, B_i(t) - \alpha' (1 - \phi)\, \psi\, A_i(t)\right] \frac{da_d}{dt} = \gamma^2 a_d . \qquad \text{(XVII.E-23)}$$

We may write this out explicitly using the solutions of $\mathbf{A}_i$ and $\mathbf{B}_i$ as

$$\frac{d^2 a_d}{dt^2} + \left[\frac{\beta' (1 - \psi)\, B_i(0)\, e^{-\Delta_i t} - \alpha' (1 - \phi)\, A_i(0)}{\zeta\, \psi\, A_i(0) - \eta\, \phi\, B_i(0)\, e^{-\Delta_i t}}\right] \frac{da_d}{dt} = \gamma^2 a_d . \qquad \text{(XVII.E-24)}$$

While this equation (and the equivalent for b_d) are likely to be difficult to solve (except for the special case of $\beta' (1 - \psi) = \eta\, \phi$ and $\alpha' (1 - \phi) = \zeta\, \psi$,) it is amenable to various approximate approaches including time averaging. Of course, we did not have to engage in all of this complicated mathematical effort, except to illustrate that sometimes such effort is worthwhile in terms of understanding and suggesting better solution methods.

References

1. Taylor, James G., "Solving Lanchester-Like Equations for 'Modern Warfare" with Variable Coefficients", Operations Research 22 , pp. 756-770, 1974.

2. Taylor, James G., and Gerald G. Brown, "Canonical Methods in the Solution of Variable Coefficient Lanchester-Type Equations of Modern Warfare", Operations Research 24, pp. 44-76, 1976.

3. Taylor, James G., and Craig Comstock, "Force-Annihilation Conditions for Variable-Coefficient Lanchester-Type Equations of Modern Warfare", Naval Research Logistics Quarterly 24, pp. 349-371, 1977.

4. Taylor, James G., "Approximate Solution (with Error Bounds) to a Nonlinear Second-order Differential Equation", Journal of the Franklin Institute 306, pp. 195-208, 1978.

5. Taylor, James G., Recent Developments in the Lanchester Theory of Combat", pp. 773-806 in K. B. Haley, **Operational Research '78, Proceedings of the Eighth IFORS International Conference on Operational Research**, North Holland Publishing Company, New York, 1979.

6. Taylor, James G., "Prediction of Zero Points of Solutions to Lanchester-Type Differential Equations for Modern Warfare", SIAM Journal of Applied Mathematics 36, pp. 438-456, 1979.

XVIII. RANGE DEPENDENT ATTRITION

A. INTRODUCTION

In the proceeding chapter, we introduced the concept of time dependent attrition rate functions. In this chapter, we introduce the more general concept of range dependent attrition. Unlike the preceding chapter which merely added the effect of time dependent attrition rate functions to already time dependent force strengths, we must take up two concepts at once: range dependent force strengths, and range dependent attrition rate functions.

Even though we have not yet dealt with the theory of attrition rate functions, in particular, Bonder Farrell Theory,[1] we have an intuitive or experiential idea that attrition rate functions should be range dependent. We know, for example, that the further away something is, the "harder" it is to see, and that the single shot probability of hit decreases (generally) with distance. Thus, we may find the concept of range dependent attrition rate functions easier to accept than time dependent (or perhaps even constant) attrition rate functions. (Indeed, we have already alluded to that basic idea in previous chapters.)

Given range dependent attrition rate functions, it therefore follows that unless the range between the two forces is changing, the resulting attrition would be the same as we have already developed for constant attrition rate coefficients in preceding chapters. It is therefore necessary that we examine range dependency in some detail.

B. RANGE AND AGGREGATION

In this section, we begin our consideration of the dynamics of force strength densities. We have earlier alluded to the general idea of force strength densities in our consideration of the effect of area of occupation on indirect fire attrition (i.e., Linear or Quadratic attrition depending on whether occupation area is constant or not.)

Let us assume that we can divide the areas of occupation, designated by Ω_A, and Ω_B, respectively for the Red and Blue forces, into small, identical area elements. The force strength in each area element, divided by the area of the element, is a force strength density. In the mathematical limit that we let these area elements go to zero, these force strength densities become actual density functions. We shall retain both of these formulations as being useful.

For the density function, we shall use the designations ρ_A, and ρ_B, for the Red and Blue forces respectively, and assume that these are functions of position $\underline{r}$ (a

vector,) time and possibly velocity $\underline{v}$ (a vector). These density functions are related to the force strengths by

$$A(t) = \int_{\Omega_A} \rho_A(\underline{r},t)\, d\underline{r}\,, \tag{XVIII.B-1}$$

and

$$B(t) = \int_{\Omega_B} \rho_B(\underline{r},t)\, d\underline{r}\,. \tag{XVIII.B-2}$$

We may define the centers of force strength (mass?) $\underline{R}_A$ and $\underline{R}_B$ of these aggregated forced strengths by

$$\underline{R}_A(t) = \int_{\Omega_A} \underline{r}\, \rho_A(\underline{r},t)\, d\underline{r}\,, \tag{XVIII.B-3}$$

for the Red force, and similarly $\underline{R}_B$ for the Blue force. Note that these centers of force strengths are time dependent. We may define the velocity of the center of force strength in the usual fashion

$$\begin{aligned} \underline{V}_A(t) &= \text{Limit}_{\Delta t \to 0} \frac{\underline{R}_A(t+\Delta t) - \underline{R}_A(t)}{\Delta t} \\ &= \frac{d}{dt}\underline{R}_A(t)\,. \end{aligned} \tag{XVIII.B-4}$$

Of course, this velocity may be different from the velocity dependence of the force strength densities.

The finite area force strength densities may be defined from the density functions. If each area element is located at $\underline{r}_{ij}$, and has area $\Delta x\, \Delta y$, then the density in area element i,j is just

$$\rho_{Aij}(t) = \frac{\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} dx \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} dy\, \rho_A(\underline{r},t)}{\Delta x\, \Delta y}\,, \tag{XVIII.B-5}$$

assuming $\underline{r}$ is two dimensional. (Extension to three dimensions at this point is left an an exercise for the student.)

The center of force strength and the total force strength in terms of these

finite area force strength densities are then just double sums over all area elements i,j with $\underline{r}_{ij}$ in Ω,

$$A(t) = \sum_{i,j \text{ in } \Omega_A} \rho_{Aij}(t), \tag{XVIII.B-6}$$

and

$$\underline{R}_A(t) = \frac{\sum_{ij \text{ in } \Omega_A} \underline{r}_{ij}\, \rho_{Aij}(t)}{A(t)} = \frac{\sum_{ij \text{ in } \Omega_A} \underline{r}_{ij}\, \rho_{Aij}(t)}{\sum_{ij \text{ in } \Omega_A} \rho_{Aij}(t)}. \tag{XVIII.B-7}$$

The vector connecting the centers of force strength, defined by

$$\underline{r}_{AB} \equiv \underline{R}_B - \underline{R}_A\,, \tag{XVIII.B-8}$$

has magnitude

$$r_{AB} = |\,\underline{r}_{AB}\,| = |\underline{R}_B - \underline{R}_A|, \tag{XVIII.B-9}$$

which is just the distance between the two centers of force strength. As we shall see, this quantity has a contral role in basic spatial aggregation. Two other vectors (and their magnitudes,) may also be useful: the position $\underline{R}^*$, of the maximum force strength density location, defined by the vector differential relationship for Red,

$$\underline{\nabla}\, \rho_A(\underline{R}^*_A, t) = 0\,, \tag{XVIII.B-10}$$

where $\underline{\nabla}$ is the vector differential operator (called a gradient when operating in this manner,) defined by

$$\underline{\nabla} \equiv x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z}\,, \tag{XVIII.B-11}$$

x, *y*, *z* are unit vectors in the x, y, z directions; and the minimum separation between the two forces, designated by

$$r'_{AB} = \mathrm{MIN}\left(|\underline{r}_A - \underline{r}_B|\right), \text{ all } \rho_A(\underline{r}_A, t)\,,\ \rho_B(\underline{r}_B, t) > 0\,, \tag{XVIII.B-12}$$

subject to the condition that the density functions of the components are non-zero.

Note that this minimum separation ceases to be meaningful when the two forces overlap (have interpenetrated,) and the position of the maximum force strength is meaningful only if the distribution function has one maximum.

It may also be useful to calculate the various moments of the distribution function. For example, the first vector moment

$$\Delta \underline{R}_A(t) \equiv \frac{1}{A(t)} \int_{\Omega_A} \left(\underline{r} - \underline{R}_A\right) \rho_A(\underline{r},t)\, d\underline{r}\,, \tag{XVIII.B-13}$$

is zero, while the second moment

$$\Delta \underline{R}_A^2(t) \equiv \frac{1}{A(t)} \int_{\Omega_A} \left(\underline{r} - \underline{R}_A\right)^2 \rho_A(\underline{r},t)\, d\underline{r}\,, \tag{XVIII.B-14}$$

is not, but is more readily understood if we write the three basic moments as

$$\begin{aligned}
\sigma^2_{Axx} &= \frac{1}{A(t)} \int_{\Omega_A} x^2\, \rho_A(\underline{R}_A + \underline{r},t)\, dx\, dy\,, \\
\sigma^2_{Ayy} &= \frac{1}{A(t)} \int_{\Omega_A} y^2\, \rho_A(\underline{R}_A + \underline{r},t)\, dx\, dy\,, \\
\sigma^2_{Axy} &= \frac{1}{A(t)} \int_{\Omega_A} xy\, \rho_A(\underline{R}_A + \underline{r},t)\, dx\, dy\,.
\end{aligned} \tag{XVIII.B-15}$$

In this case, the standard deviations σ_{Axx}, σ_{Ayy}, and σ_{Axy} (and the equivalent Blue force quantities,) reflect the physical size or extent of the forces. Moments oblong and perpendicular to the velocity of the force distribution may also be useful since units tend to orient along the direction of march.

Having defined these quantities, which as we have said, serve as an introduction to our further consideration of spatially distributed forces in a later chapter, we may now turn to consideration of range dependent aggregation. To accomplish this, we will consider two spatially distributed forces as shown in Figure (XVIII.B.1).

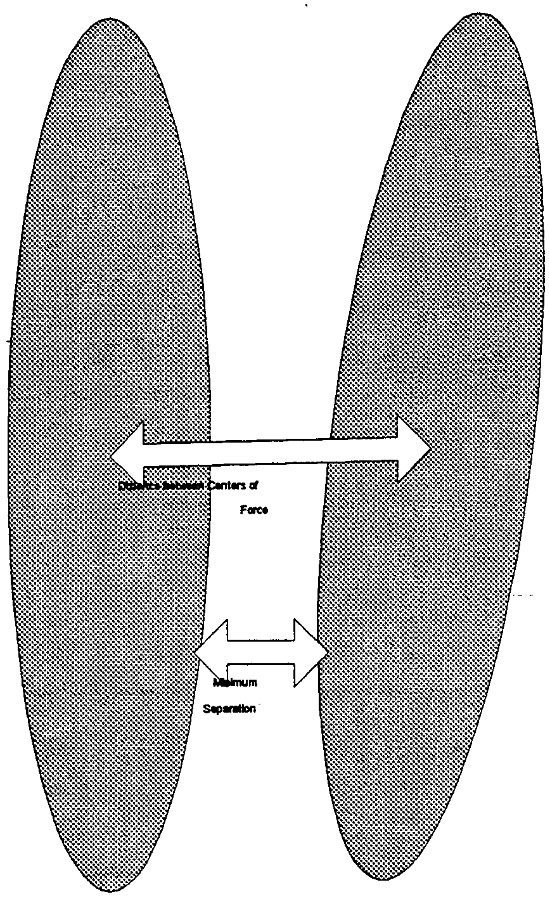


Figure XVIII.B.1 Illustrative Geometry of Two Units

so that fire is evenly distributed over all targets, then

$$\begin{aligned}\frac{d}{dt} A(t) = &-\zeta(r_{AB},t) \int d\underline{r}'_A \frac{\rho_A(\underline{R}_A + \underline{r}'_A,t)}{A(t)} \int d\underline{r}'_B \, \rho_B(\underline{R}_B + \underline{r}'_B,t) \\ &- \underline{\nabla}\zeta(r_{AB},t) \int d\underline{r}'_A \, \underline{r}'_A \frac{\rho_A(\underline{R}_A + \underline{r}'_A,t)}{A(t)} \int d\underline{r}'_B \, \rho_B(\underline{R}_B + \underline{r}'_B,t) \\ &+ \underline{\nabla}\zeta(r_{AB},t) \int d\underline{r}'_A \frac{\rho_A(\underline{R}_A + \underline{r}'_A,t)}{A(t)} \int d\underline{r}'_B \, \underline{r}'_B \, \rho_B(\underline{R}_B + \underline{r}'_B,t) .\end{aligned} \tag{XVIII.B-22}$$

We see immediately that the second and third terms on the right-hand side of this equation are zero by virtue of equation (XVIII.B-13), and the first term may be evaluated to yield a homogeneously aggregated attrition differential equation

$$\frac{d}{dt} A(\underline{R}_A,t) = -\zeta(r_{AB},t)\, B(\underline{R}_B,t) , \tag{XVIII.B-23}$$

where:

$$\begin{aligned} B(\underline{R}_B,t) &\equiv \int d\underline{r}'_B \, \rho_B(\underline{R}_B + \underline{r}'_B,t) \\ &= B(t) , \end{aligned} \tag{XVIII.B-24}$$

so that equation (XVIII.B-23) is truly the homogeneous Lanchester Quadratic attrition differential equation. Thus, we see that if we keep our honored Lanchester Assumption that fire is distributed strictly according to force strength (density), then the homogeneously aggregated Quadratic attrition differential equation results. Of course, this equation is only first order in the expansion of the attrition rate coefficient function. There are many other variations and considerations that we might retain or explore. For our purposes here, this equation will be the basis of our considerations.

If we had assumed that fire was distributed based on the actual density of targets,

$$f_B(?) = \rho_A(\underline{R}_A + \underline{r}_A,t) , \tag{XVIII.B-25}$$

and proceed as before, then the resulting first order expansion aggrgated equation is

$$\frac{d}{dt} A(\underline{R}_A,t) = -\zeta(r_{AB},t)\, A(\underline{R}_A,t)\, B(\underline{R}_B,t) , \tag{XVIII.B-26}$$

which is just a homogeneously aggregated Linear Lanchester attrition differential equation.

C. Speed and Attrition

In the homogeneously aggregated, range dependent attrition differential equation, equation (XVIII.B-22), we saw that the range dependent attrition rate function was a function of the distance between the two forces. Further, by the act of homogeneous aggregation, we reduced the spatial representation of the two forces to points. Accordingly, we may now approximately characterize the problem in terms of the scalar position equation

$$r_{AB}(t) = r_{AB}(0) - st , \qquad \text{(XVIII.C-1)}$$

where s is the speed of closure, (opening) of the two forces. This speed represents the relative speed at which the two forces are drawing together (or moving apart). This approximation is linear closure/opening at a constant rate of speed. Obviously, more general situations are possible.

Note that this equation means that the attrition rate coefficient function is (may be) an implicit function of time. We may, however, note that since the force strengths are represented by points, and therefore do not disperse, we may eliminate either time or range from equation (XVIII.B-22), writing it as either

$$\frac{d}{dt}A(t) = -\zeta(r_{AB}(0) - st)\,B(t) , \qquad \text{(XVIII.C-2)}$$

or

$$s\,\frac{d}{dr_{AB}}\,A(r_{AB}) = -\zeta(r_{AB})\,B(r_{AB}) , \qquad \text{(XVIII.C-3)}$$

although the limits of integration are different. (i.e. We integrate equation (XVIII.C-2) from 0 to t which we integrate equation (XVIII.C-3) from $r_{AB}(0)$ to $r_{AB}(t)$.)

It should also be obvious that we may apply the time dependent methods described in the proceeding chapter. For example, if we select a time increment

$$\Delta t \equiv \frac{r_{AB}(0)}{n\,s} , \qquad \text{(XVIII.C-4)}$$

assuming the range is closing, or a space increment

$$\Delta r \equiv \frac{r_{AB}(0)}{n} = s\ \Delta t\ , \tag{XVIII.C-5}$$

then with a one point (open) integral approximation, we get approximate solutions

$$\begin{aligned} A(i\Delta t) \simeq\ & A((i-1)\ \Delta t) \\ & - \zeta\left(r_{AB}(0) - (i-1)\ s\Delta t\right) B((i-1)\ \Delta t)\ \Delta t\ , \end{aligned} \tag{XVIII.C-6}$$

and

$$\begin{aligned} A\left(r_{AB}(0) - i\Delta r\right) \simeq\ & A\left(r_{AB}(0) - (i-1)\Delta r\right) \\ & - \frac{\rho_A\left(r_{AB}(0) - (i-1)s\Delta t\right)}{s} B\left(r_{AB}(0) - (i-1)s\Delta t\right) \Delta r\ , \end{aligned} \tag{XVIII.C-7}$$

for equations (XVIII.C-2) and (XVIII.C-3). If we make use of equation (XVIII.C-1) to write a parametric independent variable, either as

$$t_i \equiv i\ \Delta t\ . \tag{XVIII.C-8}$$

or as

$$r_{ABi} \equiv r_{AB}(0) - i\ \Delta r\ , \tag{XVIII.C-9}$$

the equations (XVIII.C-6) and (XVIII.C-7) reduce to

$$A_i \simeq A_{i-1} - \zeta\left(r_{AB}(0) - st_{i-1}\right) B_{i-1}\ \Delta t\ , \tag{XVIII.C-10}$$

and

$$A_i \simeq A_{i-1} - \frac{\zeta(r_{ABi-1})\ B_{i-1}\ \Delta r}{s}\ , \tag{XVIII.C-11}$$

which are identical by virtue of equation (XVIII.C-5). These equations are readily spreadsheetable. We illustrate this in Figure XVIII.C.1 where we compare the force strength trajectories for a constant, linear, and exponential attrition rates coefficient functions that have the same average value. Note the relatively slower attrition at longer range for the range dependent attrition rate coefficient function due to their smaller values at these ranges.

The other methods that we have discussed in the preceding chapter may be employed in a similar form.

At this point, it is worthwhile to consider a tactical example, originally due

Bonder and Farrell. Let the Blue force be defending and the Red force attacking. Assume that the weapons on both sides have the same functional form and dependence, and that the weapons of both sides are ineffectual at ranges greater than $r_{AB}(0)$. Since the attrition rate coefficient functions have the same functional form, a state solution, and analytical solutions, exist. These are our familiar Quadratic attrition solution.

$$A(r_{AB}) = A(r_{AB}(0)) \cosh(\tau) - \delta\, B(r_{AB}(0)) \sinh(\tau)\ ,\ r_{AB} \le r_{AB}(0)\ , \tag{XVIII.C-12}$$

and

$$B(r_{AB}) = B(r_{AB}(0)) \cosh(\tau) - \frac{A(r_{AB}(0))}{\delta} \sinh(\tau)\ ,\ r_{AB} \le r_{AB}(0)\ , \tag{XVIII.C-13}$$

where:

$$\delta = \sqrt{\frac{\zeta_A(r_{AB})}{\zeta_B(r_{AB})}} = \text{constant}\ , \tag{XVIII.C-14}$$

and

$$\tau = \frac{1}{s} \int_{r_{AB}}^{r_{AB}(0)} \sqrt{\zeta_A(r'_{AB})\ \zeta_B(r'_{AB})}\ dr'_{AB} = \int_0^t \sqrt{\zeta_A(r_{AB}(0) - st')\ \zeta_B(r_{AB}(0) - st')}\ dt'\ , \tag{XVIII.C-15}$$

with

$$t \equiv \frac{r_{AB}(0) - r_{AB}}{s}\ .$$

Note that as s increases, t decreases for any given value of r_{AB}. For the special case when the attrition rate coefficient functions are constant then

$$\tau = \sqrt{\zeta_A\ \zeta_B}\ \frac{r_{AB}(0) - r_{AB}}{s}\ . \tag{XVIII.C-17}$$

Let us now further specify the tactical situation. Blue, the defender, has a doctrine that he will withdraw when Red has closed to some range $r_{AB}^* < r_{AB}(0)$. Thus, in

fighting from $r_{AB}(0)$ to $r_{AB}{}^*$, Red takes losses.

$$\Delta A = A_0 \left(1 - \cosh(\tau')\right) - \delta B_0 \sinh(\tau') \ ,$$

where A_0 and B_0 are $A(r_{AB}(0))$ and $B(r_{AB}(0))$, and τ' is the (dimensionless) attrition time to move from $r_{AB}(0)$ to $r_{AB}{}^*$. If Red wants his losses to be as small as possible, then he must make τ' as small as possible. This means that he wants to make s as large as possible! We thus see, consistent with our Lanchestrian assumptions, that Red (and Blue,) minimizes (relatively) losses by closing (attacking) at the greatest practicable speed.

This is a completely logical conclusion that is consistent with analyses of contemporary warfare. If it takes a certain amount of time (on the average,) to find and kill a target, then the shorter the total time that targets are available, the fewer of them that can be killed. What is amazing here is that we may produce this result from what in essence is a purely Lanchestrian standpoint.

D. Conclusion and Comment

This has been a brief introduction to range dependent attrition. We could have re-elaborated the various solution techniques of the previous chapter in this range dependent attrition rate coefficient function context. However, as long as we are strictly dealing with homogeneous aggregation, range and time are equivalent, so we do not need to reexamine those techniques, but leave their translation as exercises for the student.

We must also emphasize that while we were able to derive the requisite Quadratic attrition equations by using the appropriate Lanchestrian assumption, we also made use of a very limited expansion of the attrition rate coefficient function. This represents considerable approximation in the derivation. We shall defer consideration of less (and different) approximation for a later chapter on spatially distributed forces.

REFERENCES

1. **Bonder, Seth, and Robert Farrell, eds., "Development of Analytical Models of Battalion Task Force Activities", The University of Michigan, Department of Industrial Engineering, Systems Research Laboratory, Ann Arbor, MI, Report Number SRL 1957 FR 70-1, September 1970, AD 714677.**

XIX. STOCHASTIC DUELS

A. Introduction

The subject of duels is not one that holds a familiar place in our minds. Most of our knowledge of duels derives from Hollywood, who found it a frequent dramatic vehicle in years past. Despite this, many of us have participated in one, although we may not of thought of them as such. Duels seem to be largely a male preoccupation, although there may be female versions that are not obvious to my male mind.

From an historical sense, the most famous duel that we know of as Americans is the Burr-Hamilton duel,[1] although most of us are familiar with classical duels from movies. These duels are seen as social rituals for adjudicating unacceptable behavior although the potential finality of a classical duel (to a conclusion,) may be used as a coercive or intimidating threat. In modern times, duels still exist, in the form of the automotive game of "chicken," or one-on-one basketball. Indeed, if we reduce the latter to a purely free throw competition, then the mathematics of the stochastic duel is applicable.

B. The Classical Duel

The classical duel between two participants is a strongly scripted process that (ideally) serves to settle some social disagreement between the two participants (called duelists). In contemporary fiction, the duel is portrayed as an affair of honor with the challenged party having the choice of site, time, and weapons, although social conventions may have strongly influenced all three of these. Although swords and guns were the most common weapons, one duel (in France) was (lethally) fought using billiard balls.

As we have mentioned, classical gun duels are strongly scripted. They differ from stochastic duels primarily in this manner. Thus interfiring time for classical duels is a fixture rather than a variable other than psychologically (which we shall conveniently ignore.) We limit ourselves to gun duels here, since the classical duel with edged weapons is harder to characterize mathematically, although it does bear a closer relation to stochastic duels. If we designate the individual firing probabilities of kill (or incapacitation to the point of cessation) as $p_a(n)$ and $p_b(n)$ (where the two participants or duelists are labeled a and b in our usual manner,) for the n^{th}, then after one exchange of shots (one firing,) the states are:

a dead,
b dead,
a and b dead, and
a and b alive. If we designate these states as (0,1), (1,0), (0,0), and (1,1), respectively, then the probabilities of the states are

$$
\begin{aligned}
p(0,1;1) &= p_a(1)\,\big(1 - p_b(1)\big)\,, \\
p(1,0;1) &= \big(1 - p_a(1)\big)\,p_b(1)\,, \\
p(0,0;1) &= p_a(1)\,p_b(1)\,, \\
p(1,1;1) &= \big(1 - p_a(1)\big)\big(1 - p_b(1)\big)\,,
\end{aligned}
\qquad \text{(XIX.B-1)}
$$

where the last index in p is the number of firings.

There are two mechanisms for terminating a classical duel: the death (or incapacitation) of one (or both) of the duelists, or satisfaction of the social grievance conditions that generated the duel. Neglecting the latter for our consideration, then the duel will only conclude with death. In this case, we may write the general state probabilities after n firing as

$$
\begin{aligned}
p(0,1;n) &= p_a(n)\,\big(1 - p_b(n)\big)\,p(0,1;n-1)\,, \\
p(1,0;n) &= \big(1 - p_a(n)\big)\,p_b(n)\,p(1,0;n-1)\,, \\
p(0,0;n) &= p_a(n)\,p_b(n)\,p(0,0;n-1\ ), \\
p(1,1;n) &= \big(1 - p_a(n)\big)\big(1 - p_b(n)\big)\,p(1,1;n-1)\,.
\end{aligned}
\qquad \text{(XIX.B-2)}
$$

Note that we distinguished the individual firing probabilities of kill to accommodate learning, exhaustion, and other factors. Further, only the (1,1) state is non-terminating. Since this state has the general form,

$$
p(1,1;n) = \prod_{i=1}^{n} \big(1 - p_a(i)\big)\big(1 - p_b(i)\big)\,, \qquad \text{(XIX.B-3)}
$$

we may easily perform simple calculations with these equation. In particular, we may calculate the expected number of firings before the duel is terminated as

$$
\langle n \rangle = \sum_{i=1}^{\infty} n\,p(1,1;n)\,. \qquad \text{(XIX.B-4)}
$$

Obviously, if we assume the firings to have constant probabilities, then this calculation is simplified. In this case,

$$
\begin{aligned}
p(1,1;n) &= \big(\big(1 - p_a\big)\big(1 - p_b\big)\big)^{n}, \\
&\equiv q^{\,n}\,,
\end{aligned}
\qquad \text{(XIX.B-5)}
$$

and

$$
\begin{aligned}
\langle n \rangle &= \sum_{n=1}^{\infty} n \cdot q^{n} , \\
&= \frac{q}{(1-q)^2} , \\
&= \frac{(1-p_a)(1-p_b)}{(p_a + p_b - p_a p_b)^2} .
\end{aligned}
\qquad \text{(XIX.B-6)}
$$

Similarly, we may calculate the probability of winning as

$$
\begin{aligned}
p(a) &= \sum_{n=1}^{\infty} p_a (1-p_b)\, p(1,1;n-1) , \\
&= p_a (1-p_b) \sum_{n=0}^{\infty} p(1,1;n) , \\
&= p_a (1-p_b) \sum_{n=0}^{\infty} q^{n} , \\
&= \frac{p_a (1-p_b)}{p_a + p_b - p_a p_b} ,
\end{aligned}
\qquad \text{(XIX.B-7)}
$$

and similarly for b.

Before we proceed, let us take advantage of the approximation that

$$
(1 - p_a)^n \simeq e^{-n p_a} , \qquad \text{(XIX.B-8)}
$$

and if we attribute an (average) rate of fire r to the process, then the state probabilities over time are just

$$
\begin{aligned}
p(0,1;t) &\simeq p_a\, e^{p_a}\, e^{-(p_a+p_b)rt} , \\
p(1,0;t) &\simeq p_b\, e^{p_b}\, e^{-(p_a+p_b)rt} , \\
p(0,0;t) &\simeq p_a\, p_b\, e^{p_a+p_b}\, e^{-(p_a+p_b)t} , \\
p(1,1;t) &\simeq e^{-(p_a+p_b)rt} .
\end{aligned}
\qquad \text{(XIX.B-9)}
$$

C. The Extended Classical Duel

Before proceeding to the stochastic duel, it is worthwhile to consider a classical duel extended to have multiple participants. Initially, assume the Red side

to have n_a duelists, and the Blue side m_b duelists. For simplicity, treat the firings as scripted, and the individual firing probabilities as constant.

During each firing, there are two factors that must be considered: the distribution of firings over targets - that is, some targets will be fired at more than once, while others may not be fired at; and the firings themselves - some will kill while others do not. This problem is of considerable more complexity than the previous one. As soon as n_a and m_b become very large, just enumerating the possible distribution of fire become very complicated, and as we have seen in the case of the Lanchester equations, simplify only when we have large numbers. If we assume that the extended classical duel can only end in a state (0,m), (n,0) or (0,0), then these simplifications do not exist.

It is not the intent here to fully develop this problem. Our intent was only to sketch the idea that these two degrees of freedom: target selection and kill probability; make the extended classical duel problem much more complicated than the one-on-one duel.

D. The Stochastic Duel

The stochastic duel (sometimes called the general renewal problem[2]) is similar to the classical duel except that it is not scripted. That is, firings may occur at any time. Instead of a fixed step, time dependent probability density functions $p_a(t)$, $p_b(t)$, are associated with the firing process. These functions represent the probability per unit time that a red, blue element will fire a shot. Frequently, these functions are assumed to be constant for all shots fired by each side. This is directly analogous to the assumption of $p_a(n)$ constant in the classical duel.

This probability distribution function signifies that the probability that a shot has been fired by time t is

$$P_a(1;t) = \int_0^t p_a(t')\, dt' , \qquad \text{(XIX.D-1)}$$

and the probability that two shots have been fired by time t is just

$$\begin{aligned} P_a(2;t) &= \int_{P(0)}^{P(t)} P_a(1;t - t')\, dP_a(1;t') , \\ &= \int_0^t P_a(1;t - t') \frac{dP_a(1;t')}{dt'}\, dt' , \qquad \text{(XIX.D-2)} \\ &= \int_0^t p_a(t')\, dt' \int_0^{t-t'} p_a(t'')\, dt'' , \end{aligned}$$

and so on for more shots. (The student may wish to be aware that what we are describing here is renewal theory,[3] and that it is sufficiently general to allow for different probability distribution functions for each successive shot).

If, as before in the classical duels, each firing has some probability of kill associated with it, then we may write probabilities of states (numbers n and m of each force,) at any time t. For the stochastic equivalent of the classical duel we have already developed, we have state probabilities of the form,

$$p(0,1;t) = \sum_{n=1}^{\infty} p_{ka}\,(1 - p_{ka})^{n-1}\,P_a(n;t) \left(1 - \sum_{m=1}^{\infty} p_{kb}\,(1 - p_{kb})^{m-1}\,P_b(m;t) \right), \tag{XIX.D-3}$$

which is identical in form to the equivalent equation for the classical duel. The individual terms of the first summation are the probability that after n firings, n-1 non-kills and one kill have occurred, and the probability that n firings have occurred by time t. Their product is just the probability that one kill out of n firings has occurred by time t. The first summation is thus the probability that a duelist has achieved one kill by time t.

The second summation is the same quantity for the b duelist. Thus, equation (XIX.D-3) is just the joint probability that the a duelist has achieved one kill, and the b duelist has achieved no kill, by time t. We see that this equation is fundamentally different from the equivalent classical duel equation only in the sense of being time dependent.

This equation (and the ones for the other three states,) are only a one-on-one duel. The complexities of going to the n-on-m duel are enormously greater. These complexities are beyond the scope of this text, since we are primarily and fundamentally concerned with Lanchester attrition theory. There is a considerable literature on stochastic duel theory, including both a text on one-on-one duels,[4] and numerous reports,[5] and journal articles[6]. The student who is interested in the theory and practice of stochastic duels may avail himself of these excellent sources.

E. Stochastic Duels and Lanchester Theory

To understand the relationship between stochastic duels and Lanchester attrition theory, we must delve into the rationale of stochastic duel theory a bit more in terms of how it models combat rather than its mathematical formulas. To do this, we may short circuit things by considering a theory of combat developed by C.J. Ancker, Jr. and A.V. Gafarian.

Ancker and Gafarian have been the central figures and force in the development of stochastic duel theory. Their combat theory, built in a perspective of physical phenomena and structure of theory, consists of two laws (although they might also be considered as postulates or hypothesis.[7]) The first of these is:

- *All combat is a network of firefights*;

and the second is

- *A firefight is a terminating stochastic process on a discrete state space with a continuous time parameter.*

Clearly the first law is the philosophical basis of the theory while the second is the mathematical basis.

There is a great deal of allure to and evidence for the first law of this theory. It is especially enticing to us with our highly individual view of the nature of combat. Historically, there is considerable evidence to support this view, especially with the advance of technological lethality, its resulting increase in dispersion, and the accompanying reduced size and scope of combat engagements.[a] There is also evidence against this, especially with Frederickian and Napoleonic war characterized by volley fire, but this evidence is not as compelling as the evidence for the veracity, if not the validity, of the law. Indeed, anticipating the mathematical second law for a moment, it may be constructively argued that the mathematical firefight model is sufficiently general to absorb the counter arguments as different special cases.

While we are in this historical discussion of the First Ancker-Gafarian Law, a comparison can be made with Lanchester Theory. History, and the detailed spatio-temporal analysis of battles, indicates that these battles occur as a complex of engagements between and among the subsidiary units that comprise the overall forces that are in battle. Quite frequently, the percentile losses sustained by these units may be (and are) an order of magnitude greater than those sustained by the force as a whole over the duration of the battle. An excellent example of this is the Iron Brigade at Gettysburg.

This seems to refute one of our basic assumptions of the Lanchester Attrition Theory, that fire is evenly distributed over the remaining force elements. As we have seen in sketch in other chapters, and shall deal with in detail later, this assumption is necessary to maintain the use of a rate theory when spatial aggregation is carried to homogeneity. We must recognize it therefore as fundamental to that spatial aggregation process rather than as a fundamental of the adoption of

---

[a] This leads us to an interesting question: what is the difference between an engagement and a firefight?

rate theory which is itself inherently an approximation.[b]

At a hand waving level, we know from history that units that become exhausted during a battle tend to be removed from combat and replaced with fresh units. Thus, over the course of the battle, relatively attrited units tend to be replaced with relatively unattrited units and the uniform distribution of fire tends to be a reasonable approximation under certain circumstances.

The Second Ancker-Gafarian Law, of course, complements the first law. There is little to quibble with it if we accept the first law as a definition of combat. We know that firefights are terminating and stochastic; that there are an integer number of elements involved in the firefight; and that the process occurs (non-relativistically) over time. The quibbles occur in the unique association of this law with stochastic duel theory. (And we do not believe that such a unique association is advocated by Ancker and Gafarian.)

While firefights/engagements occur in a probabilistic environment, the forces that comprise them are individuals with their own characteristics, both physical and psychological; and the forces themselves will have characteristics that embody and reflect their structure, human dynamics, military doctrine, and experience. Theory is seldom able to capture these characteristics exhaustively or completely both due to formalism limitations and political influences. These variations are almost universally reduced to a set of identical behavior probability density functions. Combat may be dominated by factors other than sheer firepower, notably terrain, maneuver, and leadership. Its termination is not well understood and is frequently represented by mathematical criteria whose rationale is based in problem simplification.

A frequent quote invoked by the proponents and develops of stochastic duels is from Clausewitz, "was is nothing but a duel on a larger scale.[8] They cite this quote as historical theoretical support for duel theory. This citation however, must be taken with a grain of salt; we may in no sense view this as some voice from the grave of Western Civilization's principal war theorist pronouncing the truth of duel theory. Taken in the context of **On War** as a whole given Clausewitz's antipathy for mathematical formalism, his stated (but not practiced,) dislike for Principles that would be reduced by practicioners to dogma, and his consistent theme of will, it seems more reasonable that it is in the latter context that the quote must be viewed.

If we return to our discussion of the classical duel, (which is the type of duel

---

[b] Regardless of mathematical approach, any theory of combat processes must admit of approximations. In a later chapter on Aggregation theory, we shall examine the common basis of modeling, the Principle of Identicality.

that Clausewitz would be familiar with,) it is the aspect of will in the termination process that Clausewitz is more probably talking about. In a classical duel, with its rigorious scripting, the participants stare death in the face up to and during each firing; only immediately after each firing do they have the option of ending the duel by some means other than death. This is a possible incentive for termination; which decision, if made on emotional basis, would, in Clausewitz's terms, be a failure of will. This is probably the exact effect that Clausewitz is talking about, that a principle factor in the termination of combat is weakening or failure of the will.

This lengthy discourse on a quotation does have a purpose. Its purpose is to illustrate that while we may logically agree that a firefight must terminate, that termination is not necessarily (nor even likely,) to be at mathematical conclusion. As we have noted previously, conclusive battles (and even engagements,) are rare; the concept of conclusion is essentially mathematical in nature and is mathematically useful. It does not generally correspond to historical evidence. We must therefore recognize that stochastic duel theoy is no different from any other mathematical attrition theory in this regard; it does not incorporate a meaningful termination mechanism (which includes breakpoints,) and we must continue to look elsewhere for such.

Nor can we disagree that combat is stochastic - even Clausewitz would agree with this. What we may quibble with is what the form of the model is. At a fundamental level, the question really has to deal with the distance that the stochastic forces operate over. The First Ancker-Gafarian Law takes an inherent view that the short range forces are more important than, even dominant over, the long range forces. While this is not an unreasonable view, being essentially the same taken in Lanchester Theory, except in weaker form, it is not the only view, notably that of Horrigan.[9]

Next, while we admit that the units engaged in the firefight have integer number of soldiers, these soldiers are not identical. Under ideal conditions, this individuality is included in the firing time probability distribution function. In combat, however, troops take injury at many levels with varying degrees of loss of effectiveness. The binary state of fully effective-killed is thus an approximation. While this approximation is common to most attrition models, including almost all formalisms of Lanchester theory, we must recognize it as an approximation.

Finally, we must recognize that as it currently stands, stochastic duel theory is spatially aggregated, just as homogeneous Lanchester theory is. This also must be viewed as an approximation.

Thus far, we have quibbled a bit with stochastic duel theory. The purpose of this is not to detract from the value of the theory; it is to demonstrate that the

theory has limitations.

One of the problems of comparing different theories is that it is difficult to sort out the differences in the meanings of the assumptions from the mathematical formalism. This is amply shown by Ancker and Gafarian in a study where they compared (homogeneous aggregated) deterministic and stochastic Lanchester theory, and stochastic duel (general renewal) theory.[10] Their conclusions are summarized as:

1. Deterministic and stochastic Lanchester mathematical formalism are not equivalent or equal.
2. For Quadratic (Deterministic) Lanchester, the force strength solutions are neither an upper nor a lower bound on the envelope of stochastic Lanchester mean force strength solutions.]
3. Even for short time, deterministic and stochastic force strength solutions may differ considerably.
4. The differences between deterministic and stochastic force strength solutions do not necessarily go to zero as $t \to 0$.
5. Stochastic (i.e. statistical) theory cannot support, in the large number limit, that individual general form firing time probability distribution fufnctions combine in an "aggregate" negative exponential firing time probability distribution function.
6. Nonhomogeneous Poisson processes, do not generally approximate general renewal processes.
7. Stochastic Lanchester and general renewal (stochastic duel) variations may have considerable magnitude, even at short times, and may (generally do) differ significantly from each other.
8. Other calculated quantities: number of survivors, battle duration, and probability of winning; differ even more among the theories than the force strength solutions.

Some of these conclusions are "old hat" to us from our previous discussions, while others are news. Thus, some comments are dictated. Before doing so, there is a viewpoint, somewhat different from the norm, that the student may find useful. Simply stated, that viewpoint is this: stochastic Lanchester theory is a special case of stochastic duel theory with delta function firing times probability distribution functions, and only one fire allocation channel. To explain these last two points a bit, delta function probability distribution functions yield fixed intervals between firings, which effectively discretize time. The concept of fire allocation channels is exactly that introduced in our earlier discussion of Lanchester-Poesson theory: that out of m firers on n targets, some of the targets will be engaged by a single firer, some by two, etc. The firing channel concept turns this around to talk about the distribution of how many of the m firers fire at one target, how many at two, etc. The single channel that is operant in stochastic Lanchester is the one where all m

firers fire at one target. (This is a natural result of the $O(\Delta t)$ reduction of state transitions.)

Stated in other words, consider the probability density matrix of stochastic Lanchester theory as our model. In stochastic Lanchester theory, any state (matrix element,) (m,n), $m \le m_0$, $n \le n_0$ is accessible from at most two status (m+1,n) and (m,n+1). In stochastic duel theory, any state (m,n) is accessible from any other state (m',n'), $m_0 \ge m' \ge m$, $n_0 \ge n' \ge n$.

Now returning to Ancker and Gafarion's findings. The first four are not surprises given our earlier discussions. They are primarily validations of the differences in the assumptions between deterministic and stochastic Lanchester. The fifth finding is important based on the argument that the firing time probability distribution function in Bonder-Farrell theory must aggregate to being negative exponential. Simply put, this finding means that one cannot derive Lanchester theory from stochastic duel theory. To proponents of stochastic duel theory, this is a death knell for Lanchester theory[c].

If we take a leaf of comparison from physics, we may see that this is not the case. If we start with General Relativity, we may derive Classical Mechanics, but not Quantum Mechanics. If we start with Quantum Mechanics, we may derive Classical Mechanics, but not General Relativity. While the reasons are different, the result is the same. We may infer stochastic Lanchester from deterministic Lanchester but not go backwards. Similarly, we may get stochastic Lanchester from stochastic duels. The fundamental reason that we cannot make either of these transformations is the dominating effect of the fundamental assumptions - the incompatibility of the mathematics merely confirms this.

This does not mean that either of the theories is inherently right or wrong universally. Each is valid under different conditions. The primary problem arises when the situation at hand requires aspects of both theories for solution and the need for a more general theory arises. A secondary problem arises when the situation at hand is such that both theories should apply.For these situations, the question is which theory is better? This implies not only accuracy but utility as well.

It is in this second situation that our inability to conduct experiments is most restrictive. This is a lack of empirical evidence to allow comparison. Thus, we may only compare the models based on their mathematical weaknesses within the framework of our model of combat, not on our observations of the actual event.

---

[c] From our standpoint, this is not the case. Lanchester attrition theory may be viewed as an application of rate theory with the stochastic attrition rate coefficient theory of Bonder-Farrell as a conjugate. This is a markedly different outlook than viewing deterministic Lanchester theory as a subset of stochastic Lanchester theory. This difference in views merely reduces Ancker and Gafarian's results to a description of difference.

(Actually the problem is further complicated by the complexity of the stochastic duel formalism. We do not expect the rate approximation to hold once m^2 or n^2 become relatively small. Since the complexity of a stochastic duel calculation is ~ m! n!, then meaningful comparisons are actually computationally limited.)

The sixth Ancker-Gafarion finding basically says that stochastic duels in the special cases that renewal theory becomes nonhomogeneous Poisson, where they are less complex, are not generally applicable. Similarly, the seventh finding is not surprising, given our previous discussions of stochostic Lanchester, and our special view of stochastic Lanchester as a special stochastic duel. Finally, we are not particularly surprised by the eighth finding: we know Lanchester losses do not relate well with history; battle duration is, in our opinion, a discredited extension of the conclusion which is only of mathematical utility; and since probability of winning is purely a stochastic concept, we would expect major differences between stochastic Lanchester and duel theories, given the channel differences.

In conclusion, we have seen that stochastic duel theory is a complicated attrition theory that mathematically implements a technically pleasing, but spatially aggregated, combat model. It subsumes stochastic Lanchester theory as a special case, but not basic or deterministic Lanchester theory, primarily due to the different enabling assumptions of the two theories. This lack of transference, plus technical difficulties with Bonder-Farrell theory are the chief basic of disagreement between the two theories since they cannot be compared directly due to the computational complexity of stochastic duel theory.

References

1. Churchill, Winston S., **History of the English Speaking Peoples**, Greenwich House, New York, 1983.. Some historians have speculated that this duel was directly responsible for preventing the young United States of America from succumbing to that degenerative disease of juvenile republics - dictatorship.

2. Ancker, C. J., Jr., and A. V. Gafarian, **Modern Combat Models: A Critique of Their Foundations**, Operations Research Society of America, 1992, and "The Validity of Assumptions Underlying Current Uses of Lanchester Attrition Rates", *Naval Research Logistics* 34 505-533, 1987..

3. Saaty, Thomas L., **Mathematical Methods of Operations Research**, Dover, New York, 1988, pp. 295-298

4. Ancker, C. J., Jr., **One-on-One Stochastic Duels**, Operations Research Society of America, 1982.

5. Gafarian, A. V., D. G. Harvey, Y. G. Hong, and M. D. Kronauer, "Some Many-on-Many Homogeneous Stochastic Combat Simulation Models", Industrial and Systems Engineering Department, University of Southern California, Los Angeles, CA 90089-0193, August 1989.

6. Gafarian, A. V., and C. J. Ancker, Jr., "The Two-on-One Stochastic Duel", *Naval Research Logistics Quarterly* **31** 309-324, C. Bernard Barfoot, "Continuous-Time Markov Duels: Theory and Application, *loc. cit.* **36** 243-253, and A. V. Gafarian, and K. R. Manion, "Some Two-on-Two Homogeneous Stochastic Combats", *loc. cit.* **36** 721-764,.

7. Ancker, C. J., Jr., and A. V. Gafarian, "An Axiom Set (Laws) for a Theory of Combat", US Army TRAC, WSMR, NM, and University of Souther California, Los Angeles, CA, Feb. 1989, and Ancker, C. J., Jr., "The Foundation for a Theory of Combat: How Modern Combat Models Fail", University of Southern California, Los Angeles, CA, August 10, 1992. The latter is included in the Proceedings of the Countermeasures Study Group led by Mr. Robert DeKinder, USA Atmospheric Sciences Laboratory.

8. Clausewitz, Carl von, **On War**, M. Howard and P. Paret, trans. and eds., Princeton, 1976.

9. Horrigan, Timothy, J., "Aggregation, Chance, and Validity in Combat Simulation", Horrigan Analytics, Chicago, 1967.

10. Ancker, C. J., Jr., and A. V. Gafarian, "The Validity of Assumptions Underlying Current Uses of Lanchester Attrition Rates", **Naval Research Logistics** 34 505-533, 1987.

XX. Stochastic Lanchester

A. Introduction

Thus far, we have largely treated the attrition problem as if it were fully deterministic. This treatment implies a fatalism about the unremitting on-slaught of attrition that is the fundamental source of repugnance for many of those who cannot accept that war (or its replacements and surrogates,) is probably a fundamental component of human nature and society. The historical data are unequivocal, back to the earliest recorded evidence,[1] while the contemporary evidence does not facilitate the confident forecast of a future that is either utopian or peaceful, despite dire predictions of the demise of warfare.[2]

In the sense of a modern world view, largely courtesy of the quantum physics developed from the start of this century,[a] and now ingrained as part of both our conscious and unconscious minds. We freely accept as temporary determinism the expectation values of sharply peaked probability distributions, or the non-occurrence of catastrophes (from the human perspective,) whose expected times to occur are at least, longer than our attention spans, and more often, longer than our lifetimes.

Just as we accept that a falling body reaches and keeps a constant velocity (if it falls far enough,) as a result of momentum loss to randomly oriented and occurring collisions with air molecules, and we describe this process "deterministically" with Newtonian mechanics, so too do we accept Lanchestrian attrition as a "smooth" mechanistic process representing the random transfer of attrition between the attrition generating elements of one force to the attrition receiving elements of the other force.[b] In this sense, then, we treat Lanchestrian attrition as a deterministic process much as we treat the effect of drag on the motion of a falling body as a deterministic process. We know that the use of deterministic drag has accuracy limits inherent to it that depend on the physical conditions. If we want greater accuracy under any set of conditions, we may have to resort to some other form of process model.

---

[a] The modern world view has also been shaped by that other great Twentieth Century physics, Relativity, but to a lesser extent. Both of these build on the advances of the previous century, Relativity on the theory of electromagnetics, and Quantum Mechanics on electromagnetics, classical mechanics, and of course, probability theory. All of these have their impact on the Physics of War as we continue to explore here.

[b] It is interesting to note that there is an analog to both Linear and Quadratic Lanchester attrition in the drag experienced by a moving body. See Marion, **Classical Dynamics of Particles and Systems**, Academic Press, New York, 1965, pp. 64-66 , whose example of this is drawn from the exterior ballistics of gun projectiles.

This imposition of accuracy limits under a given set of conditions does not compromise the use of the model implicitly nor inherently. Rather, they serve to definitely limit the degree of accuracy that we may attain by use of the model. If we cannot use the model to perform simulation to the desired degree of accuracy, then we are faced with the difficult choice of either accepting less than desired accuracy, or using a more accurate model. This does not preclude either understanding or engineering. It is instructive that the preponderance of engineering calculations performed in the design of both airplanes and missiles are performed using the classical drag model in one form or another. If we had to perform run-of-the-mill engineering aerodynamic simulation using quantum theory, air travel would likely still be limited to unpowered lighter-than-air balloons.

In a like manner, we can easily identify that the attrition rate coefficients (functions) in Lanchester theory are expectation values of probability distributions just as drag coefficients are. Just as there are conjugate theories that allow us to calculate (estimate) drag coefficients from (inherently probabilistic) quantum mechanics for use in (deterministic) classical mechanics, there are conjugate theories that provide a connection between probabilistic models of combat and (deterministic) Lanchester theory. As we have already indicated, the most compelling and useful of these conjugate theories is the Bonder-Farrell attrition rate theory that we shall take up in later chapters.[c]

Just as there are alternate expressions of drag models that take the stochastic interactions into account in different ways, there are alternate expressions of Lanchester Theory (and other combat theories) that consider the stochastic processes of combat in different ways. In this chapter, we examine other formalisms that are commonly associated with considering greater probabilistic (or stochastic) realism in the context of Lanchester theory. In one sense, these formalisms may be considered as add-ons since they come after Lanchester theory both chronologically and intellectually . In another sense, they may be thought of as more general theories since they tend to reduce to our conventional view of Lanchester theory under the appropriate conditions. In both these senses, these formalisms are (potentially) more accurate models than Lanchester theory. They are definitely sources of insight into both the mechanics and the modeling of war.

Before proceeding to take up these formalisms (or models,) it is worthwhile to present some initial insight into what we are about with this discussion of stochas-

[c] Of course, there are other theories. We have already sketched a simple such theory in the early chapters of this book where we approximated the attrition rate based on a single shot probability of kill and a firing rate. For some large scale simulations, the attrition rates may be estimated from historical experience or even from the professional opinions of experts (as in the currently popular Corps Battle Simulator or CBS simulation,) largely ignoring the impact of technology and doctrine evolution.

tic Lanchester models. Let us consider the case of a friendly force of size m units firing exactly m shots (i.e. exactly one each,) at an enemy force of size n units that we will assume for simplicity is greater in size ($n \geq m$) than the friendly force. We specify that the enemy (target) force is larger than the friendly (firing) force and that the shots are completely independent to simplify the calculation. That is, no enemy unit is shot at more than once. While we recognize that this is highly unlikely under the uncontrolled conditions we think of as combat, we adopt these specifications to make the calculations simple while retaining a flavor of the stochastic nature of combat.

If the probability of kill per shot (p) is the same for all shots (also unlikely!,) then we see immediately that the distribution of kills for all shots is binomial. (That's why we had to pose such an unlikely situation.) That is, the probability of exactly j kills out of m shots is

$$P(j) = \binom{m}{j} (1 - p)^{m-j} p^{j}. \qquad \text{(XX.A-1)}$$

We may now proceed to ask some questions: First, how many kills do we expect to occur in this bout of firing? From our previous work, we know that this is just the expectation value of j,

$$\langle j \text{ kills} \rangle = \sum_{j=0}^{m} j \, P(j), \qquad \text{(XX.A-2)}$$

where $\langle\rangle$ indicates an expectation value. We know what this result is from our earlier presentations, but we will review this calculation here in a different form, We may write equation (XX.A-2) as

$$\begin{aligned} \langle j \rangle &= \sum_{j=0}^{m} j \binom{m}{j} (1 - p)^{m-j} p^{j} \\ &= \sum_{j=0}^{m} j \binom{m}{j} q^{m-j} p^{j}, \end{aligned} \qquad \text{(XX.A-3)}$$

which serves to define q.

Equation (XX.A-3) may be rewritten as

$$\langle j \rangle = p \frac{\partial}{\partial p} \sum_{j=0}^{m} \binom{m}{j} q^{m-j} p^{j}, \tag{XX.A-4}$$

but by the Binomial Theorem, equation (XX.A-4) is just

$$\begin{aligned} \langle j \rangle &= p \frac{\partial}{\partial p} (q + p)^m \\ &= p\ m\ (q + p)^{m-1} \\ &= p\ m, \end{aligned} \tag{XX.A-5}$$

since q + p = 1 by definition.

It is also useful to calculate the variance of the number of kills. The variance is defined by

$$\begin{aligned} \sigma_{jj}^2 &= \langle (j - \langle j \rangle)^2 \rangle \\ &= \langle j^2 - 2j\langle j \rangle + \langle j \rangle^2 \rangle. \end{aligned} \tag{XX.A-6}$$

Since the expectation value is both linear and idempotent (i.e. repeated operations do not change the result,) we may further expand equation (XX.A-6) as,

$$\begin{aligned} \sigma_{jj}^2 &= \langle j^2 \rangle - 2\langle j \rangle^2 + \langle j \rangle^2 \\ &= \langle j^2 \rangle - \langle j \rangle^2, \end{aligned} \tag{XX.A-7}$$

and calculate the expectation value $\langle j^2 \rangle$ as

$$\begin{aligned} \langle j^2 \rangle &= \left(p \frac{\partial}{\partial p} \right)^2 (q + p)^m \\ &= m \left(p \frac{\partial}{\partial p} \right) p\ (q + p)^{m-1} \\ &= m\ p + m(m - 1)p^2. \end{aligned} \tag{XX.A-8}$$

The variance is then just, from equation (XX.A-7),

$$\begin{aligned} \sigma_{jj}^2 &= mp + (m^2 - m)\ p^2 - m^2 p^2 \\ &= mpq. \end{aligned} \tag{XX.A-9}$$

From these, we may now compute evolution equations for the enemy force.

The expected change in the enemy force strength n is just

$$\begin{aligned}\Delta\langle n\rangle &= -\langle j\rangle \\ &= -mp,\end{aligned} \tag{XX.A-10}$$

where we have shown an expectation of n. We do this because n may be reduced by any number between 0 and m (which we have fixed,) based on the way we have structured the firing. This is the heart of the stochastic approach - in this one-sided example, we now associate a probability distribution with the attrition process, albeit only on one side. Notably, we have assumed that the probability of kill is a singular, fixed value and is not itself stochastic. This is a common problem or limitation with most stochastic formalisms.

Obviously, since we have associated a probability distribution and a mean with n, we may describe its probability distribution in terms of higher moments. In particular, we may consider its variance,

$$\Delta\sigma_{jj}^2 = \Delta\langle j^2\rangle - \Delta\langle j\rangle^2. \tag{XX.A-11}$$

To evaluate this, we note that expectation value and finite difference operations commute (they are independent and may be performed in reversed order.) We replace squared terms with the factorial finite difference representation, yielding

$$\begin{aligned}\Delta\sigma_{nn}^2 &= \langle\Delta n^{[2]} - \Delta n^{[1]}\rangle - \Delta\langle n\rangle^{[2]} + \Delta\langle n\rangle^{[1]} \\ &= 2\langle n\Delta n\rangle - \Delta\langle n\rangle - 2\langle n\rangle\Delta\langle n\rangle + \Delta\langle n\rangle \\ &= 2\langle n\Delta n\rangle - 2\langle n\rangle\Delta\langle n\rangle.\end{aligned} \tag{XX.A-12}$$

If we now make use of equation (XX.A-10) (and its pre-expectation value equivalent,) equation (XX.A-12) becomes just

$$\begin{aligned}\Delta\sigma_{nn}^2 &= 2\langle jn\rangle - 2\langle j\rangle\langle n\rangle \\ &= 2\sigma_{nj}^2,\end{aligned} \tag{XX.A-13}$$

whose left hand side is just the covariance of n and j!

Thus far, all we have been doing is just dry statistics, but we are now poised to leap once more into attrition. Consider that some period of time Δt may be associated with the firing of these shots, or if they are repeated, with the "period" of the firing cycle. In this case, we may proceed exactly as we did in the drag example, approximating the punctuated momentum transfers of the molecular collisions as a continuous rate process. We thus replace the probability of kill with an attrition

rate times the time interval,

$$\alpha \, \Delta t \equiv p, \qquad \text{(XX.A-14)}$$

which defines α.

If we now substitute equation (XX.A-14) into equation (XX.A-10) and take the limit as $\Delta t \to 0$, we get

$$\frac{d\langle n \rangle}{dt} = -\alpha \, m, \qquad \text{(XX.A-15)}$$

and if we assume (in a rather cavalier manner,) that

$$\sigma_{nj}^2 = p \, \sigma_{nm}^2, \qquad \text{(XX.A-16)}$$

then equation (XX.A-13) may be rewritten as

$$\frac{d\sigma_{nn}^2}{dt} = 2 \, \alpha \, \sigma_{nm}^2. \qquad \text{(XX.A-17)}$$

Obviously, since m is fixed in our treatment (i,e, not stochastic,) the meaning of this last equation is problematic. On the other hand, equation (XX.A-15) is just the equivalent of our familiar Quadratic Lanchester attrition differential equation. To address this problem, we need to treat the situation in a more general manner where (at least) both force strengths are stochastic in nature.

B. Stochastic Differential Equations

Before we proceed further with our discussion of stochastic Lanchester theory, we need to spend some space considering part of a special class of differential equations, known as stochastic differential equations.[3] First, we consider a set of m differential equations. For homogeneously aggregated Lanchester theory, m is just two, but for the more complicated heterogeneous aggregation we shall consider later, m may (will) be greater in size. These differential equations have the form,

$$\frac{df_i}{dt} = b_i(t,\{f_j\}) + \sigma_{ik}(t,\{f_j\}) \, \eta_k(t) \, , \qquad \text{(XX.B-1)}$$
$$i = 1..m, \; k = 1..n,$$

where:

f_i are the random variables,
b_i are their rates of change, possibly dependent on the variables and time,
σ_{ik} are a set of functions that are related to the variances and are multiplied by the
η_k that are white noise terms which are functions only of the parametric independent variable (nominally time,) and
$\{f_j\}$ indicates the set of random variables.

The description of the $\eta_k(t)$ as white noise means that they are random functions with special properties. Specifically the white noise terms have zero expected value,

$$\langle \eta_k(t) \rangle = 0, \tag{XX.B-2}$$

and are, at once, both perfectly decorrelated and have no memory,

$$\langle \eta_k(t)\ \eta_l(t + \tau) \rangle = \delta_{kl}\ \delta(\tau). \tag{XX.B-3}$$

because the Kronnecker delta function δ_{kl} is nonzero only for $k = l$, and the Dirac delta function $\delta(\tau)$ is nonzero only for $\tau = 0$.

Since we do not normally consider probability distribution functions to necessarily have zero mean (expectation value,) we may want to think of the stochastic process depicted by equation (XX.B-1) to be decomposed into two parts: an expected value part, represented by the b_i, and a shifted (to zero mean) part that contains the higher moments, represented by the σ_{ik}.[d] The correlations between f_i and f_k are contained in the σ_{ik}. One particular aspect of this formalism of stochastic differential equations is that the random variables, the f_i, are Gaussian or normal in distribution, and stochastic processes which have this behavior are called Gaussian processes. This will be important as we proceed.

A property of these stochastic differential equations is that the probability distribution function of their random variables and time, $P(\{f_j\},t)$ obeys a differential equation known as the Fokker-Planck equation[4]

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{m} \frac{\partial}{\partial f_i} (b_i\, P) + \frac{1}{2} \sum_{i,j=1}^{m} \frac{\partial^2}{\partial f_i\, \partial f_j} (a_{ij}\, P), \tag{XX.B-4}$$

where:

---

[d] We must be careful with using this view too heavily. The b_i, as we shall see, contribute to the higher moments.

$$a_{ij} = \sum_{k=1}^{m} \sigma_{ik}\, \sigma_{kj}. \qquad \text{(XX.B-5)}$$

Under normal circumstances, the boundary conditions are specified as initial conditions for both the random variables,

$$f_i(t_0) = f_{i0}, \qquad \text{(XX.B-6)}$$

and the probability distribution function,

$$P(\{f_i\},t_0) = \prod_{i=1}^{m} \delta(f_i - f_{i0}). \qquad \text{(XX.B-7)}$$

Since this is a probability distribution function, the expected value of any function of the random variables, at any time t, is given by

$$\langle h(\{f_i\},t)\rangle = \int_{-\infty}^{\infty} h(\{f_i\},t)\; P(\{f_i\},t)\; d\{f_i\}, \qquad \text{(XX.B-8)}$$

where:

$$d\{f_i\} = \prod_{i=1}^{m} df_i, \qquad \text{(XX.B-9)}$$

is used as a shorthand and obviously indicates m integrals over all the random variables.

To illustrate the properties of this formalism, we consider an example. For the single differential equation (m = 1),

$$\frac{dx}{dt} = f(x) + g(x)\, \eta(t), \qquad \text{(XX.B-10)}$$

from which we may build the appropriate Fokker-Planck equation by inspection as

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(f\,P) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g^2\,P), \qquad \text{(XX.B-11)}$$

$$\frac{d\langle x\rangle}{dt} = \langle f(x)\rangle. \tag{XX.B-17}$$

We may also calculate the motion of the second moment of x, its variance, as (we repress the subscripts since this is a one-dimensional problem,)

$$\frac{d\sigma^2}{dt} = \frac{d\langle x^2\rangle}{dt} - \frac{d\langle x\rangle^2}{dt}. \tag{XX.B-18}$$

Since we are dealing with the random variable (or its moments) as a continuous quantity, we may apply the chain rule to the second right hand side term above,

$$\begin{aligned}\frac{d\langle x\rangle^2}{dt} &= 2\ \langle x\rangle\ \frac{d\langle x\rangle}{dt} \\ &= 2\ \langle x\rangle\ \langle f\rangle,\end{aligned} \tag{XX.B-19}$$

and the first right hand side term may be calculated in the same manner as before

$$\begin{aligned}\frac{d\langle x\rangle}{dt} &= \int_{-\infty}^{\infty} dx\ x^2\ \frac{\partial P}{\partial t} \\ &= -\int_{-\infty}^{\infty} dx\ x^2\ \frac{\partial}{\partial x}\ (fP) + \frac{1}{2}\int_{-\infty}^{\infty} dx\ x^2\ \frac{\partial^2}{\partial x^2}\ (g^2P).\end{aligned} \tag{XX.B-20}$$

We may integrate the first right hand side of equation (XX.B-20) by parts as before

$$\int_{-\infty}^{\infty} dx\ x^2\ \frac{\partial}{\partial x}\ (fP) = x^2 fP\ \Big|_{-\infty}^{\infty} - 2\int_{-\infty}^{\infty} dx\ x\ f\ P. \tag{XX.B-21}$$

If we assume th first right hand side term of the above equation away on the grounds of boundedness, then we have

$$\int_{-\infty}^{\infty} dx\ x^2\ \frac{\partial}{\partial x}\ (fP) = -2\ \langle xf\rangle. \tag{XX.B-22}$$

Now, we integrate the second right hand side integral of equation (XX.B-20) once by parts, we get

$$\int_{-\infty}^{\infty} dx\, x^2 \frac{\partial^2}{\partial x^2} (g^2 P) = x^2 \frac{\partial}{\partial x} (g^2 P) \Big|_{-\infty}^{\infty} - 2 \int_{-\infty}^{\infty} dx\, x \frac{\partial}{\partial x} (g^2 P). \tag{XX.B-23}$$

Again invoking our faithful assumption of boundedness, we may eliminate the first right hand side term of equation (XX.B-23), and integrating once more by parts, we get,

$$\int_{-\infty}^{\infty} dx\, x^2 \frac{\partial^2}{\partial x^2} (g^2 P) = 2\, x\, (g^2 P) \Big|_{-\infty}^{\infty} + 2 \int_{-\infty}^{\infty} dx\, (g^2 P). \tag{XX.B-24}$$

Invoking boundedness <u>one more time</u>, this reduces to

$$\int_{-\infty}^{\infty} dx\, x^2 \frac{\partial^2}{\partial x^2} (g^2 P) = \langle g^2(x) \rangle. \tag{XX.B-25}$$

The equation of motion of the variance then becomes

$$\frac{d\sigma^2}{dt} = 2\, \langle g^2(x) \rangle - 2\, \langle x\, f(x) \rangle + 2\, \langle x \rangle\, \langle f(x) \rangle, \tag{XX.B-26}$$

so that if we solve equation (XX.B-17) for $\langle x \rangle(t)$, then we may (in principle) solve for $\sigma^2(t)$ as well, thus specifying the time evolution of the first two moments of the random variable x.

The astute student at this point raises a good question: if we know the probability distribution function P(x,t) of the random variable x at ant time t, then why can't we just specify t explicitly and solve for the two moments? The answer is that we can, if we can get an explicit solution of P(x,t).

We conclude this section by noting that in assuming these results, equations (XX.B-17) and (XX.B-26), we have assumed away the edge effects of the distribution function. This may not always be the case. Also, this introduction to stochastic differential equations is exceedingly cursory. The diligent student may wish to pursue the subject further, using the references cited in this section or other texts.

Finally, while there are many different Fokker-Planck equations, only a few have been solved explicitly - our interest here has been to present an introduction to a formalism for analyzing attrition differential equations in a straightforward manner.

C. A Probability Matrix

In this section, we present a development method for a probability distribution function based on a finite difference basis. For this method, we consider attrition between two forces who have initial forcce strengths of m_0 and n_0, respectively, and transition probabilities per time (attrition rates) of α and β. Time is (initially) discrete winth increments Δt. The force strengths are constrained to have integer values, so that $\Delta m = \Delta n = 1$. We designate the probability density function of the force strengths m and n at any time $k\Delta t$ as p(m,n,k), recognizing that this is actually a three dimensional matrix of functions. The initial conditions on p are

$$p(m,n,0) = \delta_{m,m_0}\,\delta_{n,n_0}, \qquad \text{(XX.C-1)}$$

where we have replaced the Dirac (continuous) delta function of the preceeding section with the Kronecker (discrete) delta function as a consequence of the restriction on force strength values. We impose the additional conditions on p that

$$p(m,n,k) = 0\ ,\ m > m_0\ ,\ n > n_0, \qquad \text{(XX.C-2)}$$

and

$$p(m,n,k) = 0\ ,\ m < 0\ ,\ n < 0\ . \qquad \text{(XX.C-3)}$$

If we treat all units as firing during a time increment Δt, then we may form an evolution prescription of the form,

$$p(m,n,k+1) = p(m,n,k) + \text{Influx} - \text{Outflux}. \qquad \text{(XX.C-4)}$$

That is, the probability that the force strengths are a pair (m,n) after k+1 time increments is equal to the probability of the pair after k time increments plus an Influx of probability minus an Outflux of probability. Equation (XX.C-4) may alternately be written as

$$\Delta_k\, p(m,n,k) = \text{Influx} - \text{Outflux}. \qquad \text{(XX.C-5)}$$

in finite difference notation. We may calculate the Influx and Outflux by using the

binomial expansion technique employed in Section A of this chapter. For this derivation, we shall assume what is in essence quadratic attrition. Note that the attrition rates are taken to be fixed values, not random variables. The force strengths are the only random variables here, and they are restricted to only take on integer values.

For the Influx, we shall start by considering any force strength pair (m',n') such that m' > m, and n' > n. If, in time Δt, this pair fires exactly m' + n' times, with transition probabilities $\alpha\Delta t$ and $\beta\Delta t$, respectively, that result in exactly m' - m and n' - n kills, then the result is a force strength pair (m,n). We know from the binomial expansion that for m' firings, the probability that exactly i kills occur is

$$(\text{i kills}) = \binom{m'}{i} (1 - \alpha\Delta t)^{m'-i} (\alpha\Delta t)^{i}. \qquad \text{(XX.C-6)}$$

Since we want exactly m' - m kills to occur, we may write

$$(m'-m \text{ kills}) = \binom{m'}{m'-m} (1 - \alpha\Delta t)^{m} (\alpha\Delta t)^{m'-m}. \qquad \text{(XX.C-7)}$$

We may generalize this to write the Influx contribution from (m',n') to (m,n) as the product of probabilities of m' - m and n' - n kills and the probability that the system was in state (m',n'). Thus,

$$(m',n') \rightarrow (m,n) = \binom{m'}{m'-m}\binom{n'}{n'-n} (1 - \alpha\Delta t)^{m} (\alpha\Delta t)^{m'-m} (1 - \beta\Delta t)^{n} (\beta\Delta t)^{n'-n}\, p(m',n',k). \qquad \text{(XX.C-8)}$$

Since we know that m' can only take on values of m + 1 .. m_0 (and similarly for n'), we may write the Influx as a simple double sum

$$\text{Influx} = \sum_{i=0}^{m_0-m} \sum_{j=0}^{n_0-n} \binom{m+i}{j}\binom{n+j}{i} (1 - \alpha\Delta t)^{m+i-j} (\alpha\Delta t)^{j} (1 - \beta\Delta t)^{n+j-i} (\beta\Delta t)^{i}\, p(m'+i,n'+j,k). \qquad \text{(XX.C-9)}$$

Notice the shift of summation indices so that m + i fires by the first side kill exactly j on the second side, and the the n + j fires by the second side kill exactly i on the first side.

Calculation of the Outflux is somewhat simpler since all of the kills rise from

Calculation of the Outflux is somewhat simpler since all of the kills rise from (m,n) and we do not care in which state they end up (for the Outflux!) Thus, the Outflux term reflects the probability (m,n) effect any kills at all. This is simply the probability of any effect (= 1 by definition) minus the probability of no kills at all. For the first side (m firers,) this is just

$$1 - (1 - \alpha\Delta t)^m. \qquad \text{(XX.C-10)}$$

We must calculate the joint probabilities that no kills occur for both sides, however. It thus follows that the Outflux is simply

$$\text{Outflux} = \left(1 - (1 - \alpha\Delta t)^m (1 - \beta\Delta t)^n \right) p(m,n,k). \qquad \text{(XX.C-11)}$$

The next step that we want to take is to take the limit as $\Delta t \to 0$. Preparatory to this, we want to keep only those terms in the Influx and the Outflux that are of order Δt.

For the Influx, we see immediately, by examination of equation (XX.C-9), that since i or j must be at least 1, we need only keep the leading term in expanding the miss factors $(1 - \alpha\Delta t)$ (and its complement). Thus, the Influx reduces to two terms,

$$\begin{aligned} \text{Influx} &= \binom{m}{1}\binom{n+1}{0} \alpha\,\Delta t\; p(m,n+1,k) \\ &\quad + \binom{m+1}{0}\binom{n}{1} \beta\,\Delta t\; p(m+1,n,k) \\ &\quad + O(\Delta t^2), \end{aligned} \qquad \text{(XX.C-12)}$$

which further reduces, by the specific values of the Binomials, to

$$\text{Influx} = \alpha\; m\; \Delta t\; p(m,n+1,k) + \beta\; n\; \Delta t\; p(m+1,n,k) \qquad \text{(XX.C-13)}$$

For the Outflux, we must expand the miss factors,

$$\begin{aligned} \text{Outfux} &= \left(1 - (1 - \alpha\Delta t)^m (1 - \beta\Delta t)^n \right) p(m,n,k) \\ &\simeq \left(1 - (1 - \alpha m\Delta t)(1 - \beta n\Delta t) \right) p(m,n,k) \\ &\simeq \left(1 - (1 - \alpha m\Delta t - \beta n\Delta t) \right) p(m,n,k) \\ &\simeq \left(\alpha m\Delta t + \beta n\Delta t \right) p(m,n,k). \end{aligned} \qquad \text{(XX.C-14)}$$

$$f(x + a) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(a \frac{d}{dx} \right)^n, \tag{XX.C-19}$$

to second order (since that will be the order that we will normally calculate moments to,) to approximate the finite difference as

$$\Delta_x \simeq \frac{d}{dx} + \frac{1}{2} \frac{d^2}{dx^2}. \tag{XX.C-20}$$

This allows us to rewrite equation (XX.C-15) as a partial differential equation

$$\begin{aligned} \frac{\partial}{\partial t} p(m,n,t) = & \left(\alpha\, m \frac{\partial}{\partial n} + \beta\, n \frac{\partial}{\partial m} \right) p(m,n,t) \\ & + \frac{1}{2} \left(\alpha\, m \frac{\partial^2}{\partial n^2} + \beta\, n \frac{\partial^2}{\partial m^2} \right) p(m,n,t). \end{aligned} \tag{XX.C-21}$$

This equation replaces the matrix of probability distribution functions (of time) with a continuous probability distribution function of the random varaibles m and n (and time.)

D. A Simple Example

Having derived the evolution equations for the probability distribution function matrix, it is useful to spend some effort calculating a simple example. Let us consider the case for $m_0 = n_0 = 2$. From equation (XX.C-17), we may trivially write a solution for the initial state as

$$p(2,2,t) = e^{-2(\alpha+\beta)t}, \tag{XX.D-1}$$

which we shall see satisfies the initial conditions since $p(2,2,t) = 1$ at $t = 0$, and the other elements of the matrix will be zero at $t = 0$. We also note that if we introduce the decomposition,

$$p(m,n,t) = e^{-(\alpha m + \beta n)t}\, q(m,n,t), \tag{XX.D-2}$$

equation (XX.C-16) may be reduced to

$$\begin{aligned} \frac{d}{dt} q(m,n,t) = & \ \alpha\, m\, q(m,n+1,t)\, e^{-\beta t} \\ & + \beta\, n\, q(m+1,n,t)\, e^{-\alpha t}. \end{aligned} \tag{XX.D-3}$$

E. Evolution Equations I

While we can solve equation (XX.C-15) for the probability matrix elements of the probability distribution function, we may form evolution equations of the moments of the random variables m and n without actually performing this arduous task. Before we embark on this derivation, it is useful to review some of the mechanics of finite differences.

First, the finite difference of the product of two variables x_l and y_m is

$$\begin{aligned} \Delta x_l\, y_m &= x_{l+1}\, \Delta y_m + y_m\, \Delta x_l \\ &= y_{m+1}\, \Delta x_l + x_l\, \Delta y_m. \end{aligned} \tag{XX.E-1}$$

Further, the definite sum of a difference is

$$\sum_{l=1}^{L} \Delta x_l = x_{L+1} - x_1. \tag{XX.E-2}$$

With these, we may now proceed to examine the object of our intent.

The evolution of the expectation value of m is

$$\frac{d\langle m\rangle}{dt} = \sum_{m,n} m \frac{d}{dt} p(m,n,t). \tag{XX.E-3}$$

We may rewrite this as

$$\frac{d\langle m\rangle}{dt} = \sum_{m,n=0}^{m_0,n_0} m \left(\alpha\, m\, \Delta_n + \beta\, n\, \Delta_m \right) p(m,n,t), \tag{XX.E-4}$$

using equation (XX.C-16), and examine each term separately. Of course, this doesn't take into account the special forms of the differential-difference equations for the edges, but we will collect these terms after we have evaluated these terms.

First, we expand the right hand side of equation (XX.E-4) and consider each of the terms. Thus, the first term is

since the first and third series are identical except for the m = 0 terms which do not contribute. This equation may be rewritten as

$$\sum_{m,n=0}^{m_0,n_0} \beta\, m\, n\, \Delta_m\, p(m,n,t) = -\beta \langle n\rangle + \sum_{n=0}^{n_0} \beta\, n\, p(0,n,t), \tag{XX.E-10}$$

We may now rewrite equation (XX.E-4) as

$$\begin{aligned}\frac{d\langle m\rangle}{dt} = &-\sum_{m=0}^{m_0} \alpha\, m^2\, p(m,0,t) - \beta\,\langle n\rangle \\ &+ \beta \sum_{n=0}^{n_0} n\, p(0,n,t) - \text{edge terms}.\end{aligned} \tag{XX.E-11}$$

The edge terms arise from the special edge differential-difference equations which we review here. For the initial state:

$$\frac{d}{dt}\, p(m_0,n_0,t) = -(\alpha m_0 + \beta n_0)\, p(m_0,n_0,t). \tag{XX.E-12}$$

For the outer edges,

$$\frac{d}{dt}\, p(m,n_0,t) = -\alpha\, m\, p(m,n_0,t) + \beta\, n_0\, \Delta_m p(m,n_0,t), \tag{XX.E-13}$$

and

$$\frac{d}{dt}\, p(m_0,n,t) = \alpha\, m_0\, \Delta_n p(m_0,n,t) - \beta\, n\, p(m_0,n,t). \tag{XX.E-14}$$

And for the conclusion edges,

$$\frac{d}{dt}\, p(m,0,t) = \alpha\, m\, p(m,1,t), \tag{XX.E-15}$$

and

$$\frac{d}{dt}\, p(0,n,t) = \beta\, n\, p(1,n,t). \tag{XV.E-16}$$

The corrections that we need to use on equation (XX.E-11) only result from equa-

tions (XX.E-15) and (XX.E-16). The exclusions from equations (XX.E-12) - (XX.E-14) were for values $m > m_0$ and $n > n_0$ that we have already included by truncating the upper limits on all summations. Thus, we only need the corrections for the $m = 0$ and $n = 0$ edge terms. To derive these corrections, we rewrite equation (XX.C-16) for these edges,

$$\frac{d}{dt}\, p(m,0,t) = \alpha\ m\ \Delta_n p(m,0,t), \tag{XX.E-17}$$

and

$$\frac{d}{dt}\, p(0,n,t) = \beta\ n\ \Delta_m p(0,n,t), \tag{XX.E-18}$$

and subtract these equations from equations (XX.E-15) and (XX.E-16), respectively,

$$\delta \frac{d}{dt}\, p(m,0,t) = \alpha\ m\ p(m,0,t), \tag{XX.E-19}$$

and

$$\delta \frac{d}{dt}\, p(0,n,t) = \beta\ n\ p(0,n,t). \tag{XX.E-20}$$

We may now add these equations, multiplied by m, and summed respectively over m and n, to equations (XX.E-11). Before doing this, we note that the sum resulting from equation (XX.E-20) will be zero since $m = 0$. Thus the corrected form of equation (XX.E-11) is just

$$\begin{aligned} \frac{d\langle m\rangle}{dt} &= -\sum_{m=0}^{m_0} \alpha\ m^2\ p(m,0,t) - \beta\ \langle n\rangle \\ &\quad + \beta \sum_{n=0}^{n_0} n\ p(0,n,t) + \alpha \sum_{m=0}^{m_0} m^2\ p(m,0,t) \\ &= -\beta\ \langle n\rangle + \beta \sum_{n=0}^{n_0} n\ p(0,n,t). \end{aligned} \tag{XX.E-21}$$

This is the stochastic version of the classical Quadratic Lanchester attrition differential equation, and it incorporates the conclusion correction which can only be incorporated if we have a full solution of the proabability distribution function matrix. The other relevant moment equations, calculated in much the same manner

as we have derived equation (XX.E-21), are the evolution equation for the expectation value of n,

$$\frac{d\langle n\rangle}{dt} = -\alpha \langle m\rangle + \alpha \sum_{m=0}^{m_0} m\, p(m,0,t), \qquad \text{(XV.E-22)}$$

the variance evolution equations,

$$\begin{aligned} \frac{d\sigma^2_{mm}}{dt} &= -2\beta\sigma^2_{mn} + \beta\langle n\rangle - \beta\langle m\rangle \sum_{n=0}^{n_0} n\, p(0,n,t)\ , \\ \frac{d\sigma^2_{nn}}{dt} &= -2\alpha\sigma^2_{mn} + \alpha\langle m\rangle - \alpha\langle n\rangle \sum_{m=0}^{m_0} m\, p(m,0,t)\ , \end{aligned} \qquad \text{(XX.E-23)}$$

and the covariance evolution equation.

$$\begin{aligned} \frac{d\sigma^2_{mn}}{dt} = -\alpha\, \sigma^2_{mm} - \beta\, \sigma^2_{nn} + \alpha \sum_{m=0}^{m_0} (m^2 - m\langle m\rangle)\, p(m,0,t) \\ + \beta \sum_{n=0}^{n_0} (n^2 - n\langle n\rangle)\, p(0,n,t). \end{aligned} \qquad \text{(XX.E-24)}$$

Notice that all of these equations have conclusion corrections. From a computational standpoint these corrections represent a highly undesirable complication in that they require us to know explicitly the probability density function before we may calculate the explicit solutions of the first and second moments of the distribution. They lead us to a very natural question, "How important are the corrections?"

References

1. Farrell, Arther, **The Origins of War,** Thames and Hudson, LTD., London, 1985.

2. Crevelt, Martin van, **The Transformation of War.**

3. Hoel, Paul G., Sidney C. Port, and Charles G. Stone, **Introduction to Stochastic Processes,** Houghton-Mifflin Co., Boston, 1972.

4. Arnold, Ludwig, **Stochastic Differential Equations: Theory and Applications,** John WIley and Sons, New York, 1974.

F. The Conclusion Correction Connection

To answer this question, we have to know the explicit time dependent solution of the probability density function. We shall address this problem later in this section, both explicitly and approximately (i.e., numerically,) but first we examine the way that probability propagates through the probability density function with time. Initially, of course, only $p(m_0, n_0, 0)$ is nonzero. As time progresses, probability density spreads throughout the matrix. If we revert to the finite difference form of the probability density function evolution equation, equation (XX.C-15), then in time Δt, a kill occurs with probability α (or β) per time. Thus at each step, the total probability per time to kill one unit on the m side is $\sim \beta n < \beta n_0$. The approximate time to generate probability for killing one unit of the m side may then be approximated as

$$\delta\tau \simeq \frac{1}{\beta n_0}, \qquad \text{(XX.F-1)}$$

and therefore the total time τ for appreciable probability to reach the conclusion edge of the matrix (and thus contribute to the conclusion correction,) is approximately

$$\tau \simeq \frac{m_0}{\beta n_0}. \qquad \text{(XX.F-2)}$$

If, as is normally the case, $m_0 \sim n_0$, then for times less than approximately β^{-1}, (or alternately for $\beta t <\sim 1$,) we may ignore the correction. This is approximately the attrition time.

To see this, let us make a few simple calculations. Since probability density spread out from $p(m_0, n_0, 0)$, the conclusion edge will first be reached at state $(m_0,0)$ or $(0, n_0)$, whichever has the shorter time. Probability density reaches these states by traveling along the outer edges of the matrix. This travel has substantially simpler differential-difference equations of the form,

$$\frac{\partial}{\partial t}p(m_0,n,t) = \alpha\ m_0\ \Delta p(m_0,n,t), \qquad \text{(XX.F-3)}$$

and

$$\frac{\partial}{\partial t}p(m,n_0,t) = \beta\ n_0\ \Delta p(m,n_0,t), \qquad \text{(XX.F-4)}$$

where we have dropped the subscripts on the Δ's since there is only one force strength variable. These equations of the Poisson type that we have already solved earlier, but this solution is somewhat complcated by the solution of the initial state (m_0, n_0) evolution equation,

$$\frac{\partial}{\partial t}p(m_0,n_0,t) = -(\alpha m_0 + \beta n_0)\, p(m_0,n_0,t), \tag{XX.F-5}$$

which is

$$p(m_0,n_0,t) = e^{-(\alpha m_0+\beta n_0)t}. \tag{XX.F-6}$$

As an example, we will examine the solution of equation (XX.F-3) since we may form the solution of equation (XX.F-4) by symmetry. If we write out equation (XX.F-3),

$$\frac{\partial}{\partial t}p(m_0,n,t) = \alpha m_0 p(m_0,n+1,t) - \alpha m_0 p(m_0,n,t), \tag{XX.F-7}$$

then we may guess a solution (from our previous experience,) for $p(m_0,n,t)$ as

$$f_n\, e^{-\alpha m_0 t}. \tag{XX.F-8}$$

Substituting this into equation (XX.F-7) gives

$$\begin{aligned} e^{-\alpha m_0 t}\frac{\partial}{\partial t}f_n &- \alpha m_0 p(m_0,n,t) \\ &= \alpha m_0 p(m_0,n+1,t) - \alpha m_0 p(m_0,n,t), \end{aligned} \tag{XX.F-9}$$

which reduces to

$$\frac{\partial}{\partial t}f_n = \alpha\, m_0\, f_{n+1}. \tag{XX.F-10}$$

This is the same result that we would get for the Poisson process solution except that in this case,

$$f_{n_0} = e^{-\beta n_0 t}, \tag{XX.F-11}$$

where the initial solution for a pure Poisson process would be one.

$$p(m_0,0,t) \simeq \frac{(\alpha m_0 t)^{n_0}}{(n_0 - 1)!} \frac{1 - e^{-\beta n_0 t - n_0 + 1}}{\beta n_0 t + n_0 - 1} . \qquad \text{(XX.F-18)}$$

Since $n_0 \gg 1$, the exponential term and the 1 in the second denomiantor of equation (XX.F-18) may safely be neglected, reducing it to

$$p(m_0,0,t) \simeq \frac{(\alpha m_0 t)^{n_0}}{n_0!} \frac{1}{\beta t + 1} . \qquad \text{(XX.F-19)}$$

If we now designate the minimum probability density of interest to be $p^* < 1$, then we may rewrite equation (XX.F-19) as

$$(\alpha m_0 t)^{n_0} \simeq n_0! \, (\beta t + 1) \, p* , \qquad \text{(XX.F-20)}$$

and use Stirling's Approximation for the factorial,

$$(\alpha m_0 t)^{n_0} \simeq n_0^{n_0} \, e^{-n_0} \, (\beta t + 1) \, p* . \qquad \text{(XX.F-21)}$$

We may now take the root of both sides,

$$\alpha m_0 t \simeq n_0 \, e^{-1} \, (\beta t + 1)^{\frac{1}{n_0}} \, p*^{\frac{1}{n_0}} , \qquad \text{(XX.F-22)}$$

and since the minimum probability density raised to a very small power is essentially one, further approximate equation (XX.F-22) as

$$t \simeq \frac{n_0}{\alpha m_0} e^{\frac{\beta t}{n_0} - 1} . \qquad \text{(XX.F-23)}$$

While this equation is transcendental, it effectively proves the point.

As instructive as this demonstration has been on the mathematical impact of the conclusion edge on Stochastic Lanchester calculations, it still does not answer the fundamental question of how large are the effects. We shall address that question approximately in the next section.

G. Some Sample Calculations

If we return to the moment evolution equations of Quadratic Stochastic Lanchester, equations (XX.E-21)-(XX.E-24), we may note that the conclusion corrrections all serve to deter the rate of change of the moments. If some time is required then for probability density to accumulate in the m, n = 0 states, then a reasonable first order approximation would be to neglect the conclusion corrections when it is early in the engagement and we are far from conclusion. In this case, the evolution equations may be approximated as

$$\begin{aligned}
\frac{dA}{dt} &= -\alpha B, \\
\frac{dB}{dt} &= -\beta A, \\
\frac{d\sigma_{AA}}{dt} &= -2\alpha\sigma_{AB} + \alpha B, \\
\frac{d\sigma_{AB}}{dt} &= -\beta\sigma_{AA} - \alpha\sigma_{BB}, \\
\frac{d\sigma_{BB}}{dt} &= -2\beta\sigma_{AB} + \beta A,
\end{aligned} \tag{XX.G-1}$$

where we have returned to our usual notation for force strengths. Implicit in this approximation is the idea that force strengths (and variances and covariances,) may now take on non-integer values.

It is worth some little comment that equations (XX.G-1) possess exact solutions that may be useful. Of course, the first two equations are the familiar homogeneous Quadratic Lanchester attrition differential equations, with solutions,

$$\begin{aligned}
A(t) &= A_0\cosh(\gamma t) - \delta B_0\sinh(\gamma t), \\
B(t) &= B_0\cosh(\gamma t) - \frac{B_0}{\delta}\sinh(\gamma t),
\end{aligned} \tag{XX.G-2}$$

where: γ and δ are defined in the usual manner. The solutions for the other three differential equations are somewhat more complicated. The red-red variance is

$$\begin{aligned}
\sigma^2_{AA}(t) = {} & \frac{\sigma^2_{AA\,0}}{2}(1 + \cosh(2\gamma t)) - \sigma^2_{AB\,0}\,\delta\,\sinh(2\gamma t) \\
& - \frac{\sigma^2_{BB\,0}\,\delta^2}{2}(1 - \cosh(2\gamma t)) \\
& + A_0\left(\frac{1}{2} + \frac{\delta\,\sinh(2\gamma t)}{3} - \frac{\cosh(2\gamma t)}{6}\right. \\
& \left. - \frac{\cosh(\gamma t)}{3} - \frac{2\,\delta\,\sinh(\gamma t)}{3}\right) \\
& - B_0\left(\frac{\delta^2}{2} - \frac{\delta\,\sinh(2\gamma t)}{3} + \frac{\delta^2\cosh(2\gamma t)}{6}\right. \\
& \left. - \frac{\delta\,\sinh(\gamma t)}{3} - \frac{2\,\delta^2\cosh(\gamma t)}{3}\right).
\end{aligned} \tag{XX.G-3}$$

The blue-blue variance is similar,

$$\begin{aligned}
\sigma^2_{BB}(t) = {} & \frac{\sigma^2_{BB\,0}}{2}(1 + \cosh(2\gamma t)) - \frac{\sigma^2_{AB\,0}}{\delta}\sinh(2\gamma t) \\
& - \frac{\sigma^2_{AA\,0}}{2\,\delta^2}(1 - \cosh(2\gamma t)) \\
& - \frac{A_0}{\delta^2}\left(\frac{1}{2} - \frac{\delta\,\sinh(2\gamma t)}{3} + \frac{\cosh(2\gamma t)}{6}\right. \\
& \left. - \frac{2\cosh(\gamma t)}{3} - \frac{\delta\,\sinh(\gamma t)}{3}\right) \\
& + B_0\left(\frac{1}{2} + \frac{\sinh(2\gamma t)}{3\,\delta} - \frac{\cosh(2\gamma t)}{6}\right. \\
& \left. - \frac{2\sinh(\gamma t)}{3\,\delta} - \frac{\cosh(\gamma t)}{3}\right).
\end{aligned} \tag{XX.G-4}$$

The red-blue covariance is

$$\begin{aligned}
\sigma^2_{AB}(t) = {} & \sigma^2_{AB\,0}\cosh(2\gamma t) - \frac{\sigma^2_{AA\,0}\sinh(2\gamma t)}{2\,\delta} \\
& - \frac{\sigma^2_{BB\,0}\,\delta\,\sinh(2\gamma t)}{2} \\
& + \frac{A_0}{6\,\delta}(\sinh(2\gamma t) - 2\,\delta\cosh(2\gamma t) \\
& + 2\,\delta\cosh(\gamma t) - 2\sinh(\gamma t)) \\
& - \frac{B_0}{6}(2\cosh(2\gamma t) - \delta\,\sinh(2\gamma t) \\
& - 2\cosh(\gamma t) + 2\,\delta\,\sinh(\gamma t)).
\end{aligned} \tag{XX.G-5}$$

In common usage, the initial conditions on the variances/covariance are zero so that the three preceeding equations only have terms in the initial force strengths.

It is possible to perform calculations using these equations, or they may be approximated further as finite difference solutions (as we have done before.) To demonstrate how these functions vary, we present two sample calculations in Figures (XX.G.1), force strengths, and (XX.G.2), variances and covariances, for two cases. Both have the same initial conditions, but different attrition rate coefficients. The two cases differ in their α's only. The first is a non-draw where the α is half what it would be for a draw. The second is a draw.

There should be no surprises in the force strengths since we have seen their like many times before. The plot of the variances/covariances is new information. By examinig thesecond figure, we may make several observations. First, we note that since the variances are driven by the enemy's inflicted attrition (i.e., friendly losses,) then the red force has the larger variances. Second, the blue variances do not differ appreciably, a result that we would expect from the common β in the two cases. Last, we note that the draw case red variance and the covariance grow faster than the non-draw case - again, what we would expect from the differences in α.

Despite this, the variances grow relatively slowly when compared to sizes of the force strength. This leads us to another approximation. Since we have a handy way of approximately computing the first and second moments of the Stochastic Lanchester probability density function, it is a simple step to approximate that function with a normal distribution. Using the notation of the above equations, and m, n for the random force strength variables, this approximate probability density function has the form,

$$p_g(m,n,t) = \frac{e^{-\frac{G}{2}}}{2\,\pi\,\sigma_{AA}\,\sigma_{BB}\sqrt{1-\rho^2}}, \qquad \text{(XX.G-6)}$$

where:

$$G \equiv \frac{1}{1-\rho^2}\left(\frac{(m-A)^2}{\sigma_{AA}^2} + \frac{(n-B)^2}{\sigma_{BB}^2} - 2\,\rho\,\frac{(m-A)\,(n-B)}{\sigma_{AA}\,\sigma_{BB}}\right), \qquad \text{(XX.G-7)}$$

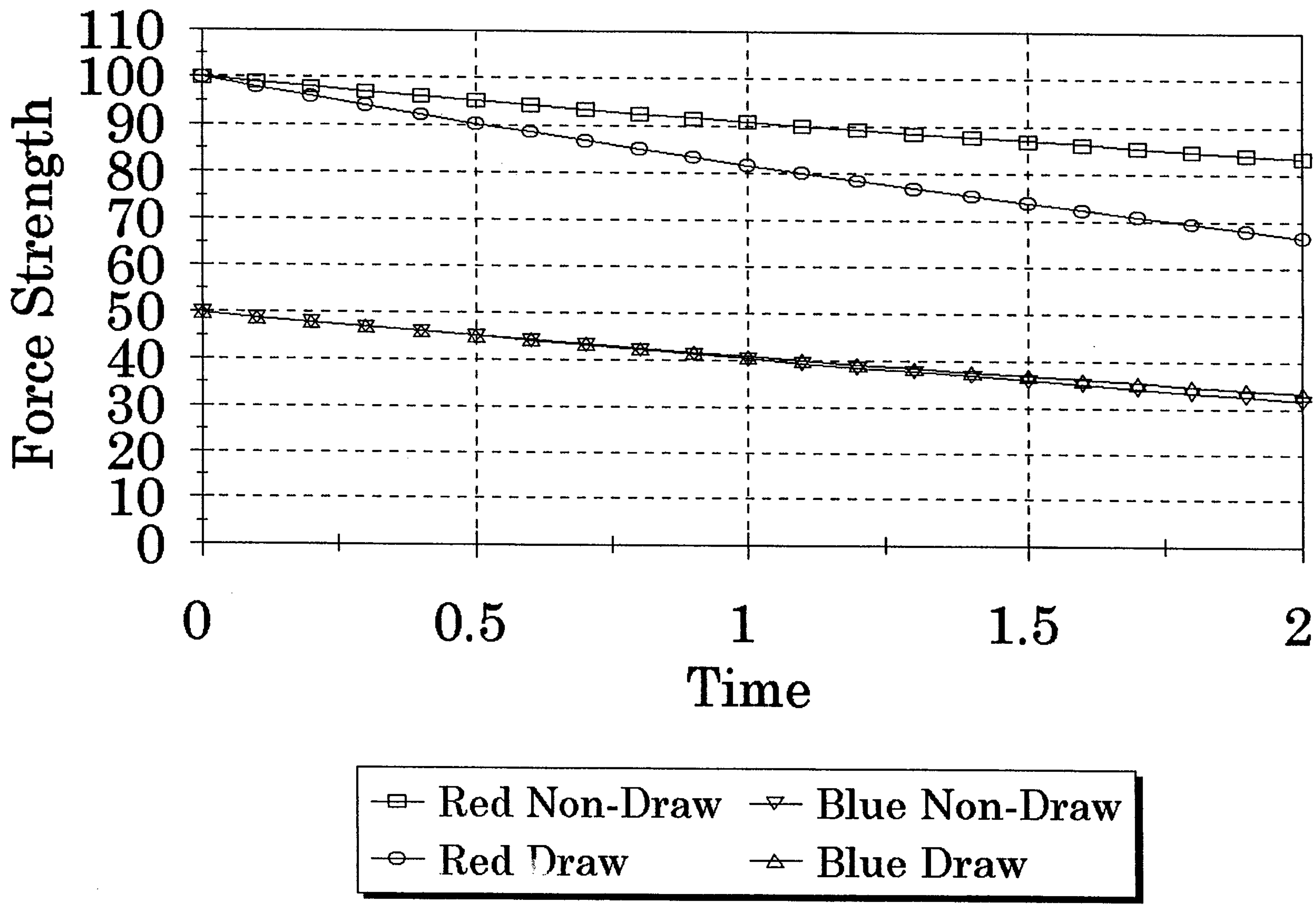


Figure XX.G.1. Force Strength Examples

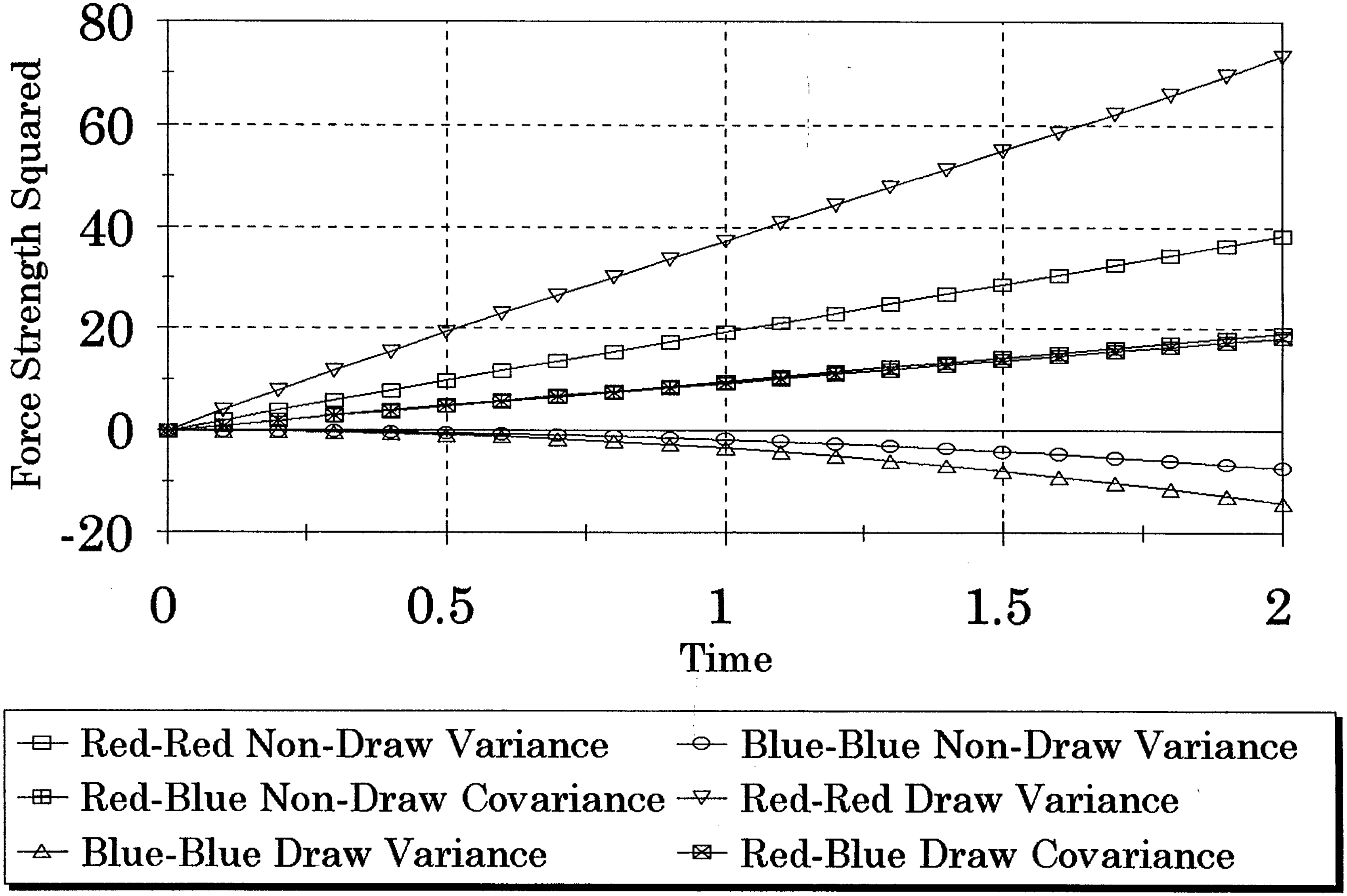


Figure XX.G.2. Variances and Covariances Examples

and the correlation coefficient is defined by

$$\rho \equiv \frac{\sigma^2_{AB}}{\sqrt{\sigma^2_{AA}\ \sigma^2_{BB}}} . \qquad \text{(XX.G-8)}$$

Using the non-draw case of the earlier example, Figures (XX.G.3)-(XX.G.7) present a representation of this approximation.

The first plot at t = 0.1, is presented rather than a t = 0 plot because the latter is simply a delta function and would not graph well. If we examine the plots in time sequence, it is immediately obvious that the peak of distribution moves with time (since A and B change with time and the distribution is peaked around them.) Further, if we examine the scales of the plots, we see that not only are the distributions widening (due to the growth of the variances and covariance,) but that the magnitude of the peak is decreasing. This is not a surprising event if we consider that we have required probability (*in toto*) to be conserved.

This is exactly the phenomena we were concerned about in the preceding section. As the probability distribution function evolves, probability density spreads out to lower force strength value states. (Technically, to make a more valid comparison, we should swithch back to an integer representation by integrating equation (XX.G-6) over each bi-unit intervals. This is, however, a reasonable approximation for visualization pruposes since the area of integration would be unity and the probability density functions varys only slowly over that area. Thus the approximate integral is just the value of the probability density function times the area of integration (= 1!).)

Before proceeding with our discussion, we need to recall that the moments calculated with equations (XX.G-1) yield values that are larger than the values that would be calculated with conclusion correction. Despite this, we see that by t = 2, by which time the red force has taked about 20% casualties and the blue force 40%, which we know from our historical examination to be extreme losses indeed, there is still no appreciable probability density in the region of the conclusion edge despite the faster transport of probability density outward. Of course, the normal approximation does not accumulate probability in the conclusion edge, but spreads it beyond into negative force strength states. To accomodate this difference, we may approximate the conclusion edge terms by

$$p(m,0,t) \simeq \int_{-\infty}^{0} dn\ p_g(m,n,t) , \qquad \text{(XX.G-9)}$$

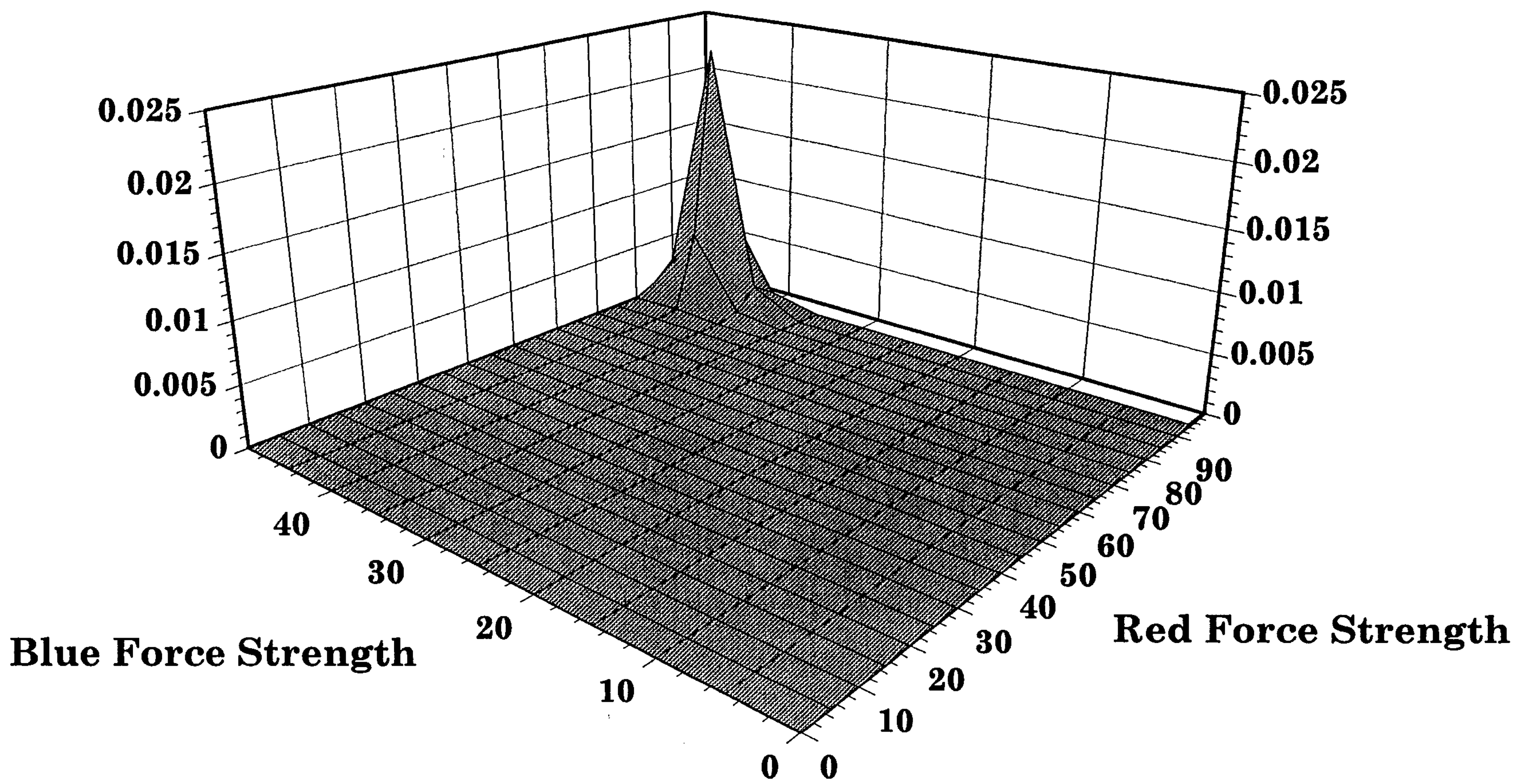


Figure XX.G.4. Gaussian Approximation Probability Density, t = 0.5

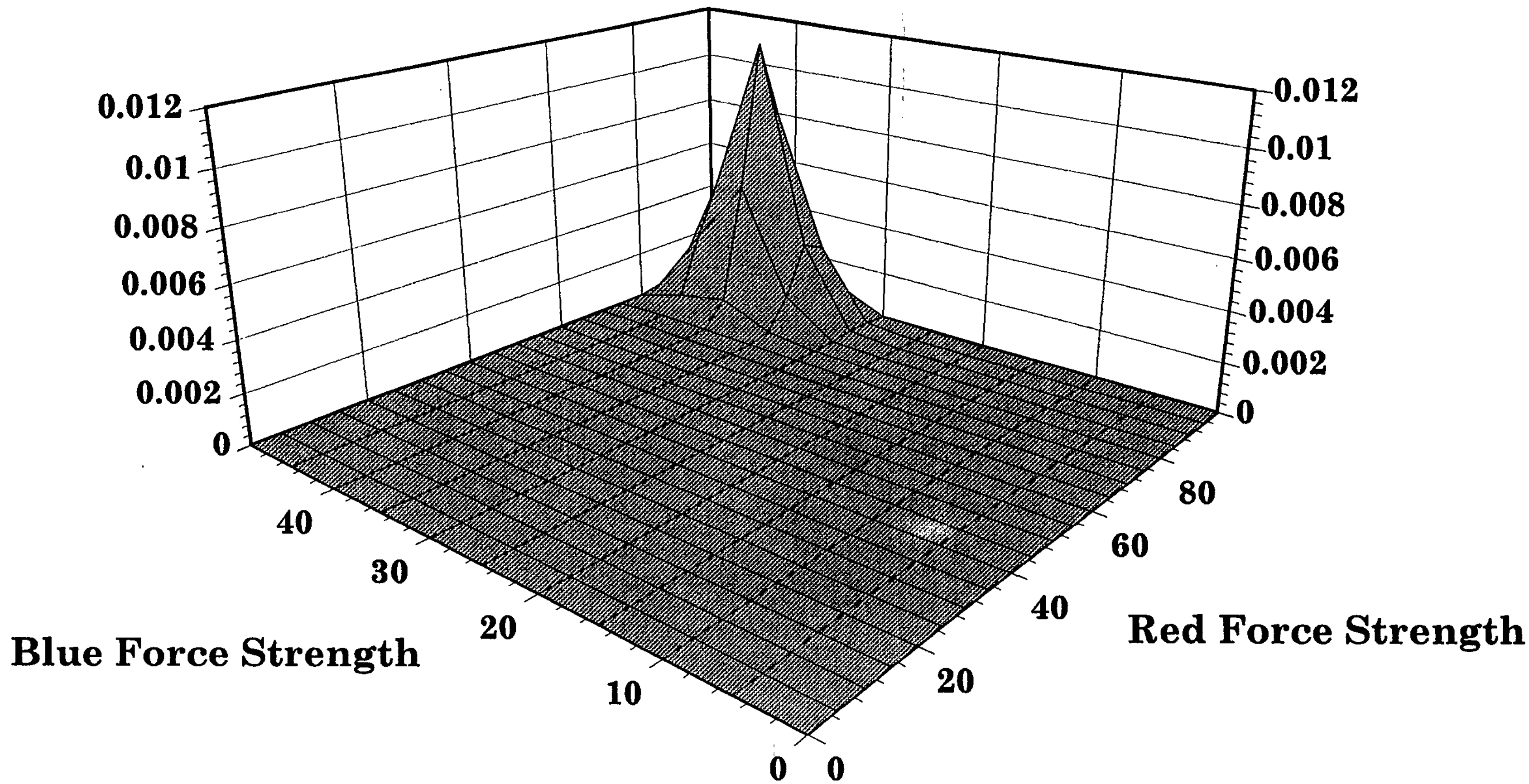


Figure XX.G.5. Gaussian Approximation Probability Density, t = 1.0

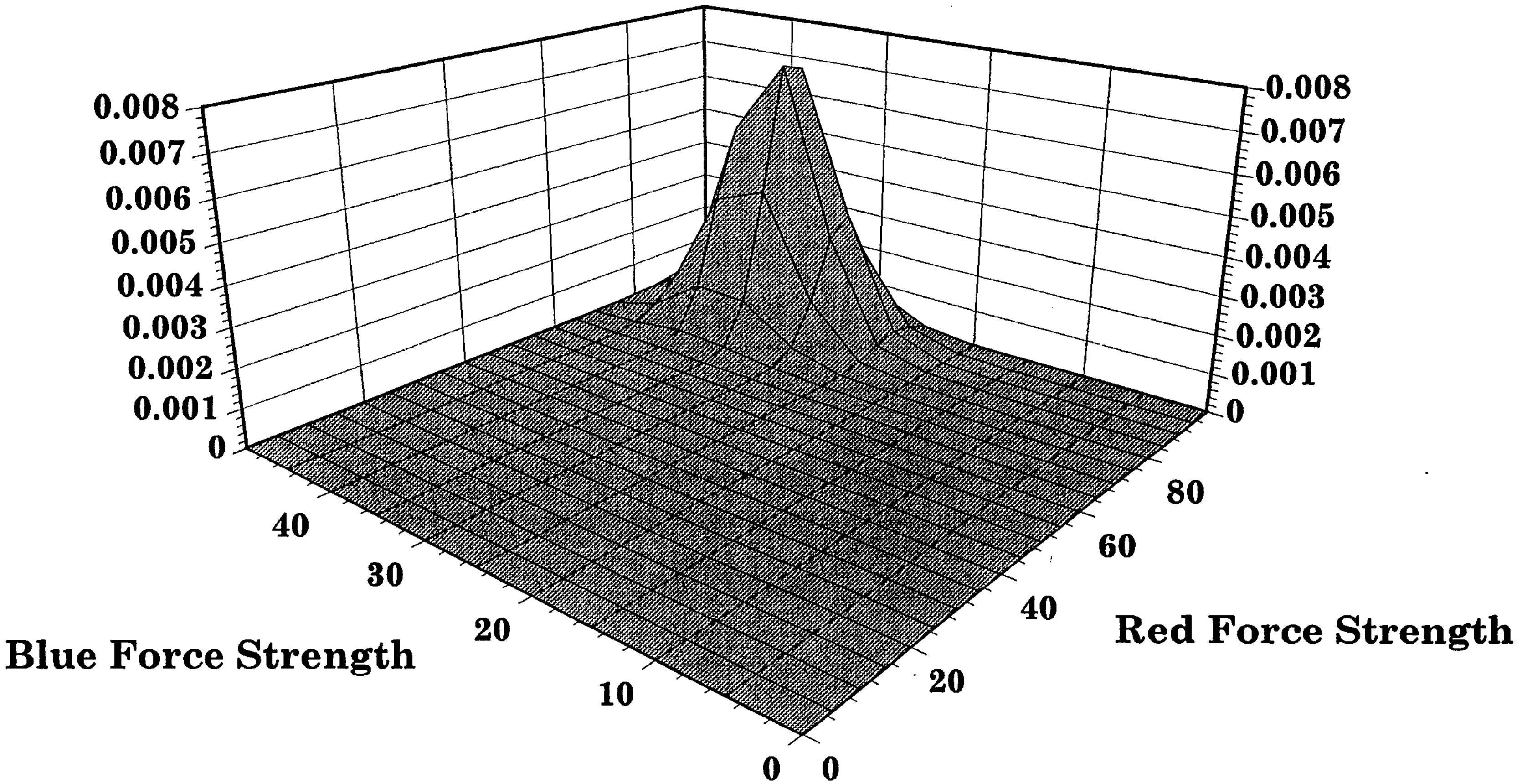


Figure XX.G.6. Gaussian Approximation Probability Density, t = 1.5

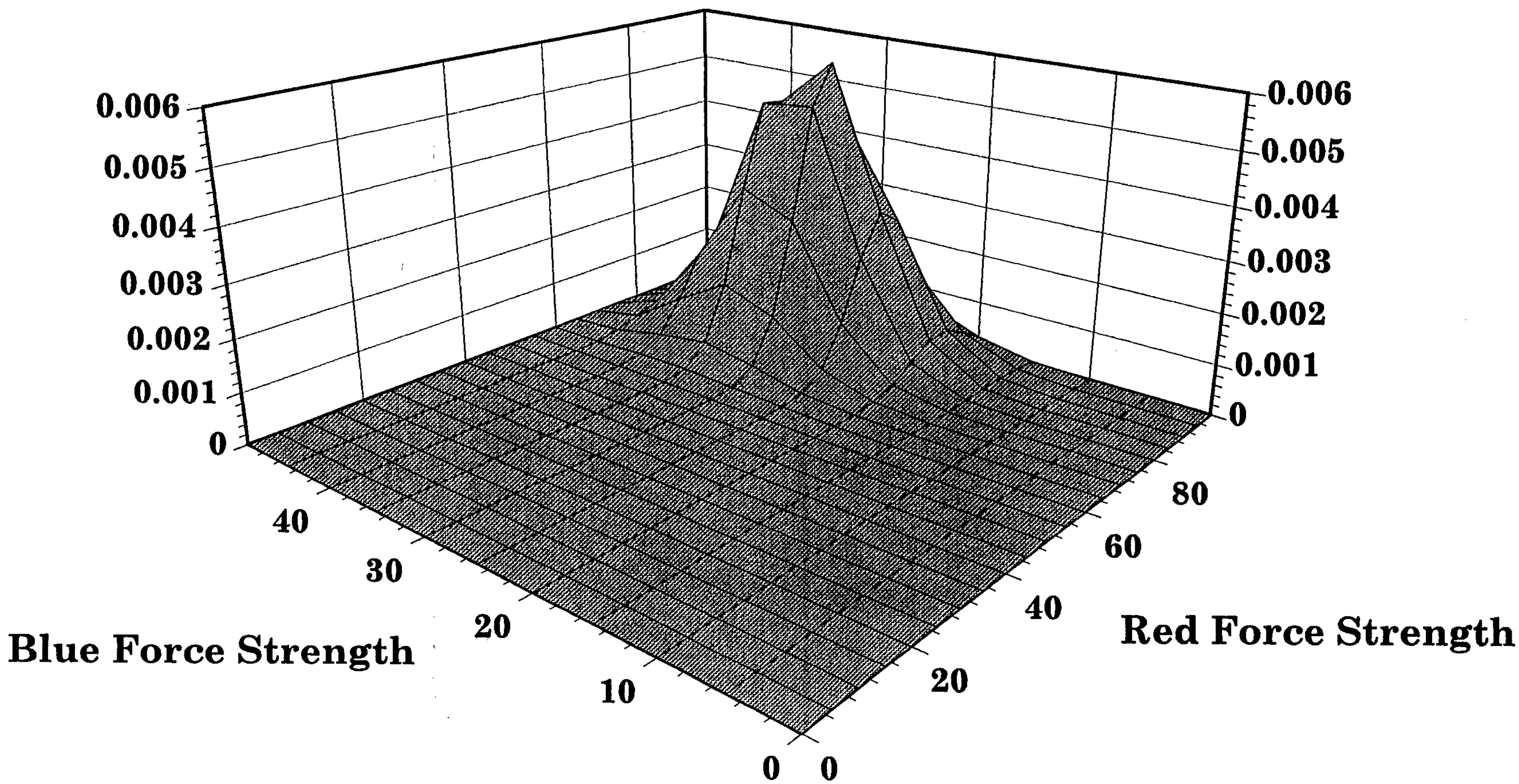


Figure XX.G.7. Gaussian Approximation Probability Density, t = 2.0

which, if we ignore the cross-correlation term as small, and apply the approximations in Appendix F, allow us to approximate the conclusion edge contribution as

$$p(m,0,t) \approx \frac{1}{\sqrt{2\pi}\,\sigma_{AA}} e^{-\frac{(m-A)^2}{2\sigma_{AA}^2}} \frac{1-\sqrt{1-e^{-\frac{2B^2}{\pi\sigma_{BB}^2}}}}{2} . \qquad \text{(XX.G-10)}$$

A similar equation may be written for p(0,n,t). From these approximations, it is then straightforward to return to the moment equations of section E and render approximations for them. For example, equations (XX.E-22) for the expected value of n is

$$\frac{d\langle n\rangle}{dt} = -\alpha\langle m\rangle + \alpha \sum_{m=0}^{m_0} m\, p(m,0,t) . \qquad \text{(XX.G-11)}$$

Since the clear intent of the conclusion correction term is to represent the number of m that survive to conclusion, we may substitute equation (XX.G-10) and replace the sum with an integral,

$$\frac{d\langle n\rangle}{dt} = -\alpha\langle m\rangle + \frac{\alpha}{\sqrt{2\pi}\,\sigma_{mm}} \frac{1-\sqrt{1-e^{-\frac{2\langle n\rangle^2}{\pi\sigma_{nn}^2}}}}{2} \int_0^{m_0} dm\, m\, e^{-\frac{(m-\langle m\rangle)^2}{2\sigma_{mm}^2}} . \qquad \text{(XX.G-12)}$$

The integral of equation (XX.G-12) can be integrated, partly exactly and partly approximately using the same techniques, to give

$$\frac{d\langle n\rangle}{dt} = -\alpha\langle m\rangle + \alpha \frac{1-\sqrt{1-e^{-\frac{2\langle n\rangle^2}{\pi\sigma_{nn}^2}}}}{2} \left(\frac{\sigma_{mm}}{\sqrt{2\pi}} \left[e^{-\frac{(m_0-\langle m\rangle)^2}{2\sigma_{mm}^2}} - e^{-\frac{\langle m\rangle^2}{2\sigma_{mm}^2}} \right] + \frac{\langle m\rangle}{2} \left[\sqrt{1-e^{-\frac{2(m_0-\langle m\rangle)^2}{\pi\sigma_{mm}^2}}} - \sqrt{1-e^{-\frac{\langle m\rangle^2}{2\sigma_{mm}^2}}} \right] \right) . \qquad \text{(XX.G-13)}$$

This is a very(!) Non-linear equation that will probably have to be solved numerically. Further, so long as $\langle n \rangle$ is large compared to σ_{nn}, the effect of the correction will be small.

H. The Stochastic Engagement

As we have already noted, the common usage is to use zero initial values for the variances/covariance. This does not necessarily need to be the case, but reflect the choice of a delta function for the boundary condition for the probability distribution function. Given the continuity of the equations, it is straightforward to have non-zero initial conditions. What do these mean? They could represent uncertainty in the strengths of the two forces at the start of the engagement.

What do we do with them? While studies have shown no correlation with strength and victory that is consistent, most strategists believe that a key factor in victory is the will of the commander and his belief in the possibility of victory. In this case, it is easy to postulate a model based on the perception of chance of victory. If the conclusion criterion of deterministic Lanchester is used as a model, then red may perceive that victory is possible if

$$A > \delta B \ . \tag{XX.H-1}$$

The probability of this criterion being satisfied is then just

$$\begin{aligned} P_A(t) &= \sum_{n=0}^{n_0} \sum_{m=\delta n}^{m_0} p(m,n,t) \\ &\simeq \int_{-\infty}^{\infty} dn \int_{\delta n}^{\infty} dm \, p_g(m,n,t) \ , \end{aligned} \tag{XX.H-2}$$

where we have (I hope obviously,) mixed the integer and non-integer uses of m and n. We cannot expect this quantity to be very useful since it will not cross over value from > (<) 0.5 to < (>) 0.5 over time (lacking reinforcements, which we have not examined.)

A more likely approach seems to be to carry two sets of variance/covariance equations, one for each commander. Then each will have a separate probability of victory based on his perceptions of his forces, the enemy's, and the engagement.

References

5. Hildebrand, Francis B., **Methods of Applied Mathematics**, Prentice-Hall, INC., Englewood Cliffs, New Jersey, 1965, pp. 224-225.

XXI. Theory of Rates

A. Introduction

Take quantities of water and table salt. Prepare a saturated solution. That is, dump salt into the water gradually, stirring vigorously to dissolve the salt. Once salt will no longer dissolve, wait a few moments for true equilibrium.

Now transfer most of the solution into another container, taking care to leave behind the residue. Into this vessel, carefully introduce a small salt crystal, leaving the vessel open to the air, wait. Over time, the seed crystal will grow.

If we view this system at the molecular level, there are three distinct regions: the crystal, the salt water solution, and the air above the solution. Each has different properties.

The crystal is an ordered structure of sodium and chloride ions, microscopically arranged in a lattice, but macroscopically possibly irregular in shape, although usually we grow salt crystals to display macroscopic symmetry that demonstrates the microscopic order.

The salt water solution is much less ordered. It consists of water molecules and sodium and chloride ions. All of these are in motion. (The ions in the crystal are in motion too, but this motion is so small we may ignore it for the purpose of our example.) Some of the water molecules are weakly attached to the ions, forming a cloud around them, while others are not attached. The speed of the unattached water molecules, and the water-ion clouds are distributed in a functional form that is determined by their mass and the temperature of the solution. Because they are moving, they bump into each other. Order is very short range, basically limited to the clouds which are themselves changing as a result of collisions.

The air above the solution is essentially not ordered at all. It consists of air molecules (oxygen, nitrogen, carbon dioxide, etc.) and water molecules. These are also in motion and are continually colliding as well. Because the density of these molecules is much less than in the liquid, these collisions are less frequent just because these are fewer of them to collide and their average distance apart is much greater.)

Most of the interesting stuff occurs on or near the interfaces of the regions. Water molecules and ion-water clouds collide with the surface of the crystal. If the water molecules hit the crystal just right, then they may pull an ion off the crystal, forming a new ion-water cloud. If an ion-water cloud hits the crystal just right, it leaves the ion on the crystal and frees up the water molecules.

Similarly, water molecules (and water-ion clouds,) may break through the surface barrier and come into the air. Likewise water and air molecules can also break through this surface barrier and enter the solution. Thus there are mechanisms for transport among the three regions.

If the entire system were in equilibrium then all of this transport would not be evident. If the solution remains saturated, then each ion that is deposited on the crystal is balanced by another that is removed. (This is simply the definition of a saturated solution.) If the air above the solution was saturated with water, then there would similarly be no net movement across the interface.

Usually, this is not the case. The air above the solution is not saturated with water. Since it is not, more water molecules enter the air than leave it. Thus the amount of water in the solution decreases and the salt concentration increases. Because the salt concentration increases more water molecules are taken up in ion-water clouds. Because there are more ion-water clouds, more ions are deposited on the crystal, and it grows. All of this occurs at the microscopic level. All of this transport occurs randomly, stochastically.

Nonetheless, we may also view this overall process at the macroscopic level. The number of collisions of water molecules and ion-water clouds with the surface of the crystal is a function of the salt concentration in the solution and the surface area of the crystal. The surface area of the crystal (if it is regular), is a function of the mass of the crystal. Thus, the rate of change of the mass of the crystal is a function of the mass of the crystal and the salt concentration in the solution. Above some concentration, the mass increases, below, it decreases.

Similarly, the rate that water mass leaves the solution (at constant temperature,) depends on the amount of water in the air. Thus the rate of change of water mass from the solution depends on an evaporation rate. Obviously, salt mass is removed from or added to the solution as it is added to or removed from the crystal. The rate of change of concentration of the salt solution is thus a function of evaporation (rate) and salt mass (change).

In principle then, we may write a set of rate equations (differential equations with respect to time,) for the crystal mass (which is a function of concentration,) and the concentration that in effect describe the time dependent mass of the crystal. Over regions where the basic assumption hold, these macroscopic rate equation provide quite accurate representations of the macroscopic effects of these microscopic phenomena.

B. Rate Process

To understand the nature of rate processes, we first return to the basic facts implicit in measurement. Although we cannot ever perform such an experiment (except in the context of a simulation or, possibly, a training exercise) we shall begin with a combat experiment.

Let us begin with two forces (Red and Blue,) that have initial force strengths A_0 and B_0, respectively. At time t_0, combat begins. At subsequent times t_1, t_2, ..., Red loses, respectively, one element, while at times, t_1', t_2', ..., Blue loses, respectively, one element. Mathematically, the Red force strength has the form

$$A(t) = A_0 - j \ , \ t_{j-1} \le t < t_j \ . \qquad \text{(XXI.B-1)}$$

The representative shape of this trajectory is shown in Figure XXI.B.1. Interestingly, it follows from this that this trajectory does not possesses a derivative with respect to time. That is, the derivative of A(t) is zero except at $t = t_j$ where the derivative is infinite. On this basis, it is difficult to see how we may describe these data in terms of a rate.

Having imagined that we have this set of data, let us now further imagine that we have N sets of such data. We further assume that N is sufficiently large and the sets were measured in such a way that these set are uniquely representative of the combat processes being studied. If we use the frequentist view of probability, then this set of trajectories may be thought of as defining a probability density function (in the summation rather than the integral sense.)

If we label the individual trajectories as $A_i(t)$ (and $B_i(t)$,) and their individual transition times as $t_{i,j}$ and $t_{i,j}'$, then the probability at time t that a force of initial strength A_0, opposing an enemy force of initial strength B_0, under the combat conditions of these data, has strength A is just

$$p(A,t) = \frac{1}{N}\sum_{i=1}^{N} \delta(A - A_i(t)) \ . \qquad \text{(XXI.B-2)}$$

Note that since the $A_i(t)$ only have integer value, the value of A must also be integer.

It is a simple matter to use equation (XXI.B-2) to calculate the expected value of A at time t as

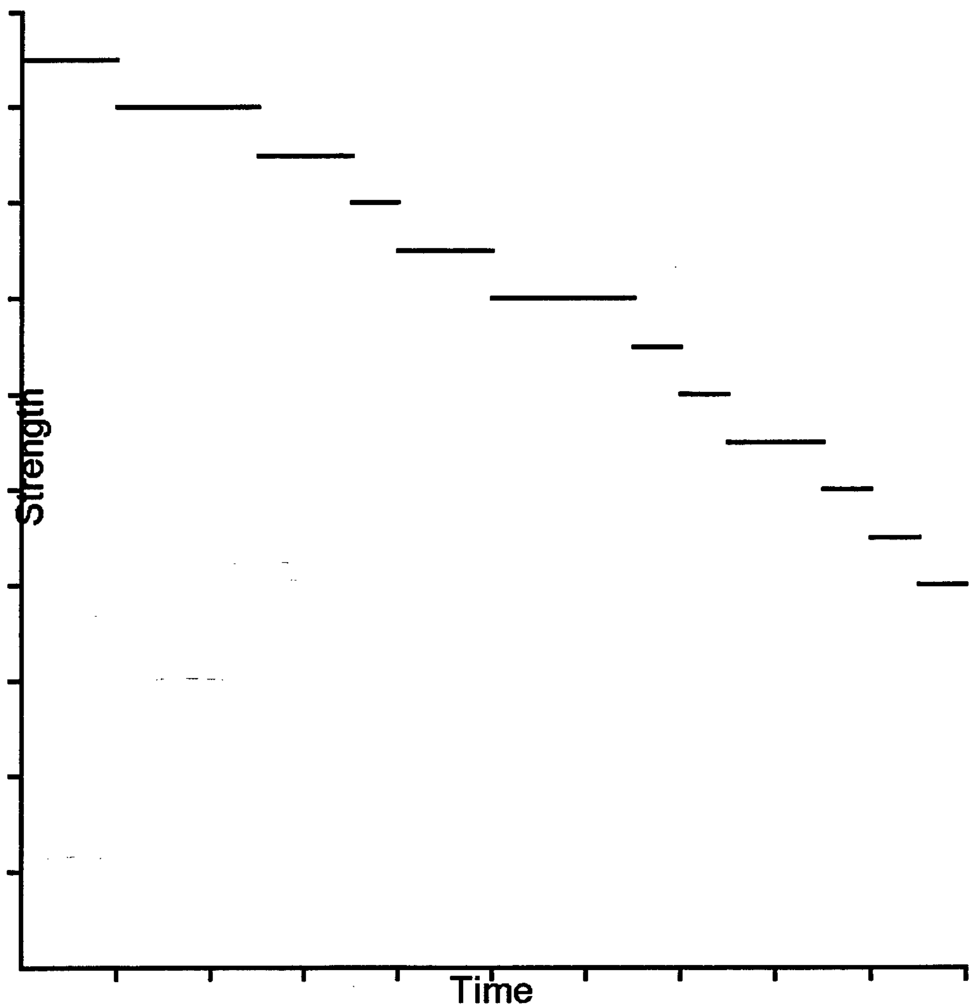


Figure XXI.B.1. Trajectory of Force Strength Experiment

$$\langle A(t) \rangle = \sum_{a=0}^{A_0} A\, p(A,t) \ . \tag{XXI.B-3}$$

We note that while the A are integer valued, $\langle A(t) \rangle$ is not, in general, integer valued.

This approach is essentially that taken in the stochastic (general renewal) duel - calculate the time dependent probability that the force is in a particular state (A), and from that probability calculate the trajectories of the moments.

The rate theory approach is considerably different from this. Each trajectory of the data set goes from value A_0 - j to A_0 - (j + 1). The average transition time is then just

$$\tau_j = \frac{1}{N} \sum_{i=1}^{N} t_{i,j} \ . \tag{XXI.B-4}$$

Thus, instead of an average trajectory of the form

$$\langle A(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} A_i(t) \ , \tag{XXI.B-5}$$

we may consider an average trajectory that takes on integer values A_0 - j at t = τ_j. Further, we may extend the definition of the trajectory to time values between the τ_j as

$$\underset{\sim}{A}(t) = A_0 - j - \frac{t - \tau_j}{\tau_{j+1} - \tau_j} \ , \ \tau_j \le t \le \tau_{j+1} \ . \tag{XXI.B-6}$$

Note that this mathematical representation is defined on a closed interval, unlike equation (XXI.B-1), which is defined on a partly open interval, since

$$\underset{\sim}{A}(\tau_{j+1}) = A_0 - (j + 1) \ . \tag{XXI.B-7}$$

Further, equation (XXI.B-6) has a time derivation,

$$\frac{d}{dt} \underset{\sim}{A}(t) = - \frac{1}{\tau_{j+1} - \tau_k} \ , \tag{XXI.B-8}$$

which defines the rate on the interval [τ_j, τ_{j+1}]. (We ignore here the fact that we

believe the rate of force strength change to be equal to some function of the force strengths times a "constant." Remember that our purpose was to establish the basis for rate theory. That we have done. Establishing the functional form of the rate is more a matter of data analysis.)

This then is the basis of rate theory. It defines an average trajectory based on a rate of change that is the inverse of the average of the time for a transition to occur. We may contrast this with the stochastic duel approach that defines an average trajectory as the instantaneous expectation of all possible states.

We may explore this difference in some more detail. To do this, we first examine the nature of our data in probabilistic terms. As the student will recall, the basic premise of stochastic duel theory is to take the continuous time probability distribution function of the interfiring time (i.e., time between shots,) and the probability of kill given a shot to calculate the joint probability distribution of the force strength states. The moment trajectories may then be computed from this joint probability distribution function.

At the detail level, there are two random events in the firing cycle of an individual fires. At the beginning of the cycle, there is a random time to complete the firing. The value and frequency of this time is determined by the interfiring time probability density function. At the end of the firing cycle, there is a chance of success given by the probability of kill given a shot. In stochastic duel theory, the mathematical prescription is to combine these probabilities interactively, allowing for losses of fires, to form the joint probability distribution function.

Let us now use this process to develop an algorithm that we can use to develop a simulation with. Assume that we have the interfiring time probability distribution functions for the two sides, Red and Blue, and the associated probabilities of kill given a shot. We designate these as $p_A(t)$ and $p_B(t)$, and $p_{k|s|A}$ and $p_{k|s|B}$, respectively. Further, we will assume that the interfiring time probability distribution functions are integrable

$$
\begin{aligned}
P_A(t) &= \int_0^t p_A(t')\, dt' , \\
P_B(t) &= \int_0^t p_B(t')\, dt' ,
\end{aligned}
\tag{XXI.B-9}
$$

and that the resulting functions are invertible,

$$
\begin{aligned}
t_A &= P_A^{-1}(r_A) , \\
t_B &= P_A^{-1}(r_B) ,
\end{aligned}
\tag{XXI.B-10}
$$

where r_A and r_B are random numbers defined on the unit interval. Equations (XXI.B-10) will allow us to calculate interfiring times explicitly.[a]

Next,we assume the two forces to be initially comprised of A_0 and B_0 elements, respectively, that are identical but distinguishable. Distinguishability is an artifice of the algorithm to allow us to track the elements of the two forces. If we assume that $t_0 = t_0'$, that is, that both forces begin combat at the same time, then we start the algorithm by randomly selecting a target from the opposing force for each element of both forces. Next, we randomly select a firing time for each element of both forces. We sort these triads of firing element, target element, and event time from least to greatest firing time. The initial event times are the first firing times.

The rest of the algorithm is to proceed through the list of triads until a firing time exceeds some preselected time value, then we stop. Each time we treat a triad in the list, we perform the following actions:

- we randomly determine whether a kill occurs using the values of probability of kill given a shot;
- if a kill does not occur, we randomly select a new firing time using equation (XXI.B-10) and add this to the previous event time to be the new firing time;
- if a kill does occur, then we remove the triad for the target element from the list, randomly select new targets for all elements that were firing at that target, generate new event times for these elements by randomly selecting firing times and adding these to the current event time[b] and note the two total force strengths and the event time;
- we sort the list again;
- and repeat the process.

This algorithm is depicted in Figure XXI.B.2.

What this algorithm, which we can readily translate into a computer simulation, tells us is the times that one of the two forces loses an element. In other words, this algorithm produces exactly the kind of information that we have been considering in this section.

Let us consider two cases of what we may expect this data to look like, one special and the other general. These are sketched in Figures XXI.B.3. and XXI.B.4.

---

[a] Actually we do not have to be able to do this analytically, but this is easier and neater than doing it numerically.

[b] Alternately, we could wait till those event times occur and then select new targets, etc. This effectively delays the rate of the combat. Which approach is selected depends on whether we model the firer as being continuously aware of the target's state or not.

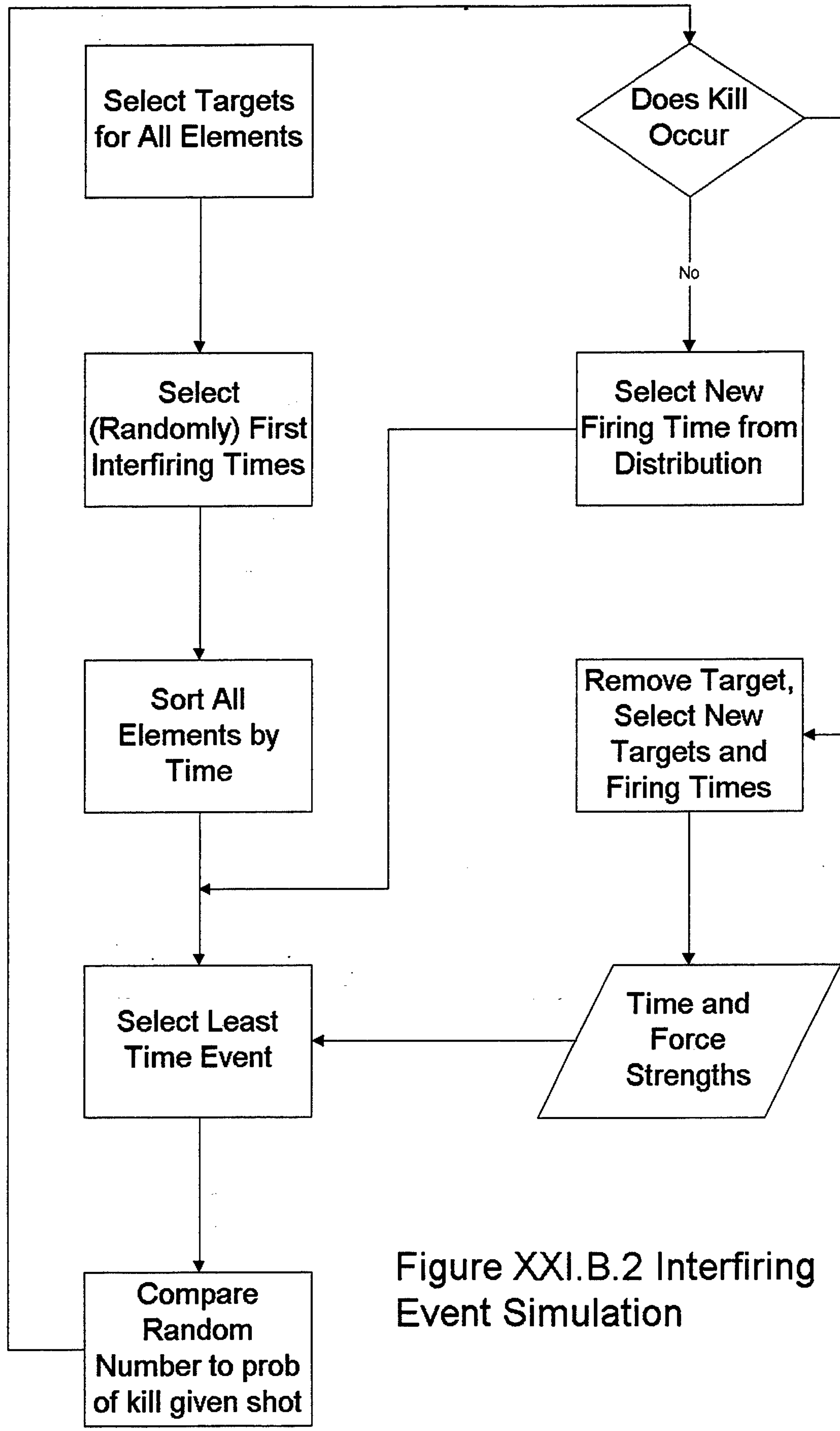


Figure XXI.B.2 Interfiring Event Simulation

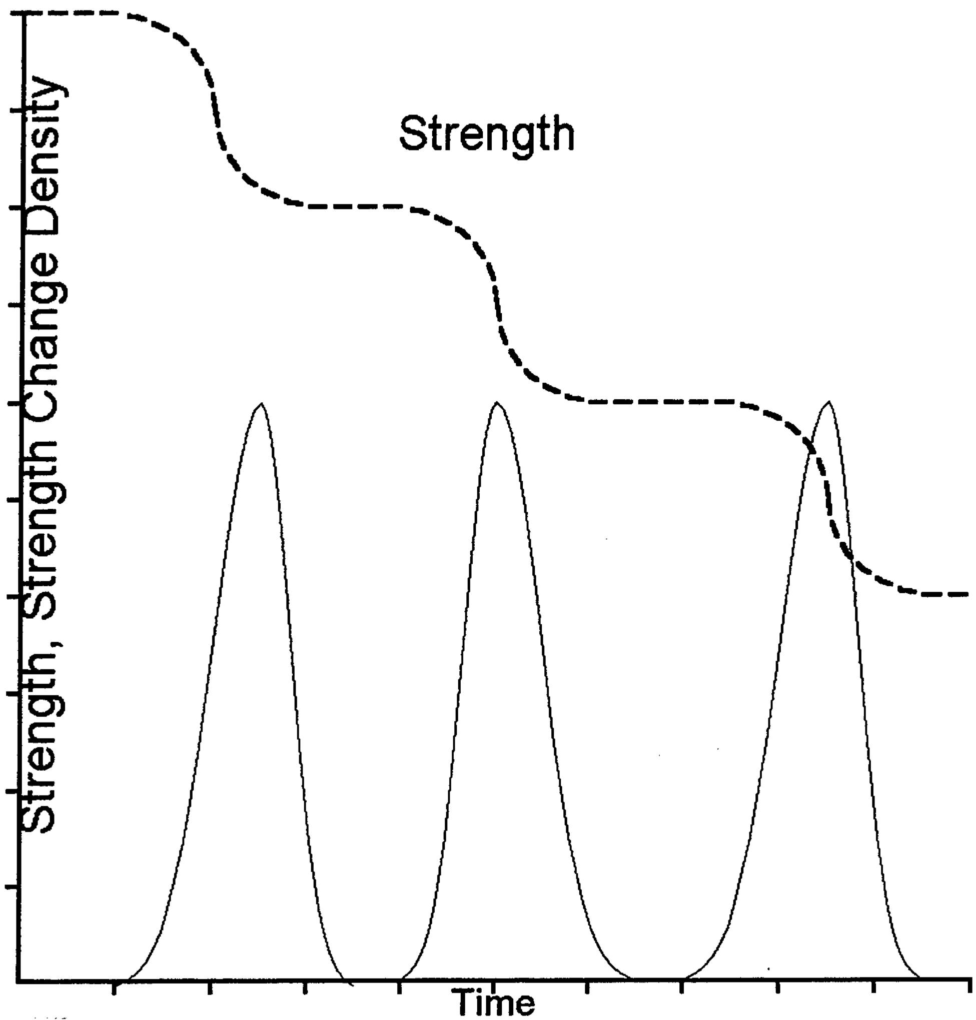


Figure XXI.B.3. Small Deviation Transitions

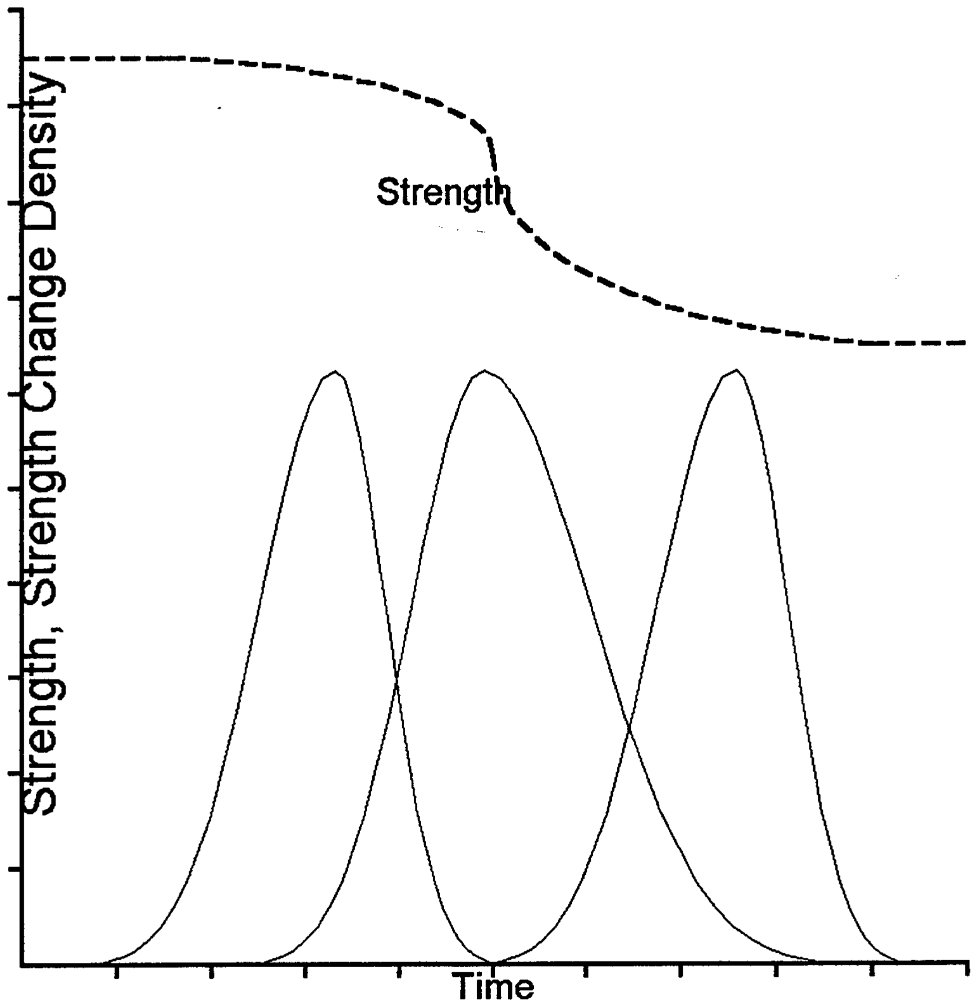


Figure XXI.B.4. Large Deviation Transitions

For the special case, let us assume that the interfiring time probability distribution functions are such that the mean or expected interfiring time is large compared to the standard deviation. That is

$$\langle t_a \rangle \gg \sigma_A \ , \qquad \text{(XXI.B-11)}$$

and similarly for Blue, where:

$$\langle t_A \rangle = \int_0^\infty t' \, p_A(t') \, dt' \ , \qquad \text{(XXI.B-12)}$$

and

$$\sigma_A^2 = \int_0^\infty (t' - \langle t_A \rangle)^2 \, p_A(t') \, dt' \ . \qquad \text{(XXI.B-13)}$$

We further assume that the probabilities of kill given a shot are close to one in value.

In this case, event times will tend to be clumped. That is, there will be periods of time when there are no or few kills, interspersed with times when there are kills. This is the situation shown descriptively in Figure XXI.B.3. The solid line depicts the density of event times, and the dashed line depicts the force strength trajectory. The detail of the latter does not accurately depict the step behavior we showed in Figure XXI.B.1, so we must imagine that it is present.

As we examine this figure, we may note several things. First, the detailed or microscopic step behavior that we noted in Figure XXI.B.1. is replicated at the gross or macroscopic level. Second, because of this macroscopic step behavior, we would expect that a rate theory approximation of this data would not be very good, but that a stochastic duel representation would. Of course, we could approximate the macroscopic steps with a rate theory approach on a piece-wise basis and have a reasonable approximation, but we would have to know when this condition would occur for us to do so. Luckily, this can be easily estimated mathematically by considering equation (XXI.B-10) and the value of the probability of kill given a shot.

Is this special case realistic? When may we expect it to occur? I will advance two cases where this seems possible, if not likely. If the forces are ground forces, and are firing in volley, then doctrine forces the interfiring probability distribution function to artificially be a delta function. In this case, the standard deviation is effectively zero. The mathematics that describes volley fire takes on a special form that we shall consider in more detail later.

The other case is essentially the same, but would be naval forces firing by

broadside. In this case, the effect is the same, but the military situations and terminology are different.

Let us now consider what is probably the more common situation. In this case,

$$\langle t_a \rangle \sim \sigma_A \, , \tag{XXI.B-14}$$

and the probability of kill given a shot is not essentially one. Now subsequent firings wthl shorter firing times, begin to overlay with firings of longer firing times, and the individual distributions overlay. Now the force strength trajectory, as shown in Figure XXI.B.4. is "smoother," and we would at once expect a rate theory approximation to be better, and a stochastic duel treatment, while still good, to be more complicated.

Before proceeding to address some simulated examples, there is one more rate approximation approach that we may consider. Instead of forming a rate based on the average valued of the inter-event times, we may take a different approach. To do this, we must change our viewpoint a bit.

To start with, we take the force strength trajectory, which at the data level are just pairs of force strength values, $A_0 - j$, and event times, $t_{i,j}$. If we treat these literally as data pairs, then we may form a rate approximation as

$$A_i(t) \simeq A_0 - j - \frac{t - t_{i,j}}{t_{i,j+1} - t_{i,j}} \, , \; t_{i,j} \le t \le t_{i,j+1} \, . \tag{XXI.B-15}$$

If all the inter-event time intervals, $t_{i,j+1} - t_{i,j}$ are approximately equal, then we may form a total interval form of this equation as

$$\overline{A_i}(t) \simeq A_0 - J \, \frac{t - t_{i,0}}{\Delta t_i} \, , \tag{XXI.B-16}$$

where

$$\Delta t_i \equiv t_{i,J} - t_{i,0} \, , \tag{XXI.B-17}$$

or

$$\overline{A_i}(t) \simeq A_0 - \frac{t}{\Delta t_i'} \, , \tag{XXI.B-18}$$

where

$$\Delta t_i' = \frac{1}{J} \sum_{k=0}^{J-1} (t_{i,k+1} - t_{i,k}) . \qquad \text{(XXI.B-19)}$$

(We note that $\Delta t = J\,\Delta t'$.)

Next, we may pointwise average the rate approximations in the usual manner, as

$$\langle A(t) \rangle \simeq \frac{1}{N} \sum_{i=1}^{N} \overline{A_i}(t) . \qquad \text{(XXI.B-20)}$$

By equation (XXI.B-15), this is just

$$\langle A(t) \rangle \simeq A_0 - \frac{1}{N} \sum_{i=1}^{N} \left[j + \frac{t - t_{i,j}}{t_{i,j+1} - t_{i,j}} \right], \qquad \text{(XXI.B-21)}$$

where we must select the proper values of $t_{i,j}$, $t_{i,j+1}$ and j for each value of t. As we have already noted, if the inter-event times are approximately equal for any particular set of data and assuming $t_{i,0} = 0$, then by using equation (XXI.B-18), we have the approximation

$$\langle A(t) \rangle \simeq A_0 - \frac{t}{N} \sum_{i=1}^{N} \frac{1}{\Delta t_i'} . \qquad \text{(XXI.B-22)}$$

This is a very interesting equation when we compare it with equation (XXI.B-8). In the first case, the rate of change of the (approximate) force strength trajectory is the inverse of the average inter-event time. In the second case, the rate of change is the average of the inverse of the inter-event time.

Despite the fact that we use the same words in describing these two rates, they are fundamentally different. Which is the better representation, the better approximation? To address this question, we shall take up the actual analysis of simulated data (since we do not have any access to real data,) in the next section.

C. Simulated Analysis

In the preceding section, we discussed an algorithm for generating simulated

force trajectory data in terms of how we might develop an approximation of this data in terms of rate theory. We deliberately ignored two factors in that discussion. First, if we actually expect this data to be approximated by Lanchester attrition differential equations, then the rate of change of the data will not be represented by a simple rate constant. Instead, the rate will be represented by (we expect,) a constant times the opposing force strength.

Second, we have not addressed the question of how we would actually perform this analysis. This is the method in our madness. We want to keep separate the basis for rate theory approximation of the data from the mechanics of actually performing the analysis of the data. Too often, we become involved in the latter and its instrumentality to the point that we believe that the theory of the instrumentality is the basic theory itself. Thus, the separation between the basic theory (in the previous section,) and the analysis in this section).

As we noted in the previous section, it is relatively easy to turn the algorithm described there into a simple computer simulation. It is then equally simple to use this simulation to generate a set of simulated data.

In fact, that is exactly what we did for three different probability density functions: a uniform probability density function, a negative exponential probability density function (NED), and a "gamma" probability density function with the specific form,

$$p(t) = \alpha^2 \, t \, e^{-\alpha t} ; \qquad \text{(XXI.C-1)}$$

for the interfiring time. In each case, the probability density function was constrained to produce an expected firing time of 2 (time units). The uniform distribution has lower and upper limits of 1 and 3, respectively. These distributions are compared in Figure XXI.C.1.

Using each of these probability density functions in turn, 25 sets of force strength trajectory data were generated for initial Red, Blue force strengths of 50 and 35, respectively, and probabilities of kill given a shot of 0.1 and 0.07, respectively. Selection of these values is arbitrary except that (a.) we want enough data sets for a reasonable sample, (b.) we want the force strengths to be large enough for a reasonable number of events per trajectory, and (c.) we don't want a draw situation.

These data provide a basis for analysis in terms of the ideas presented in the preceding section. (Because of the extensiveness of these data sets,we do not present these data in the sense of Figure XXI.B.1.)

We may, however, present averaged force strength trajectories for each of these forces and probability density functions in the manner of equation (XXI.B-3).

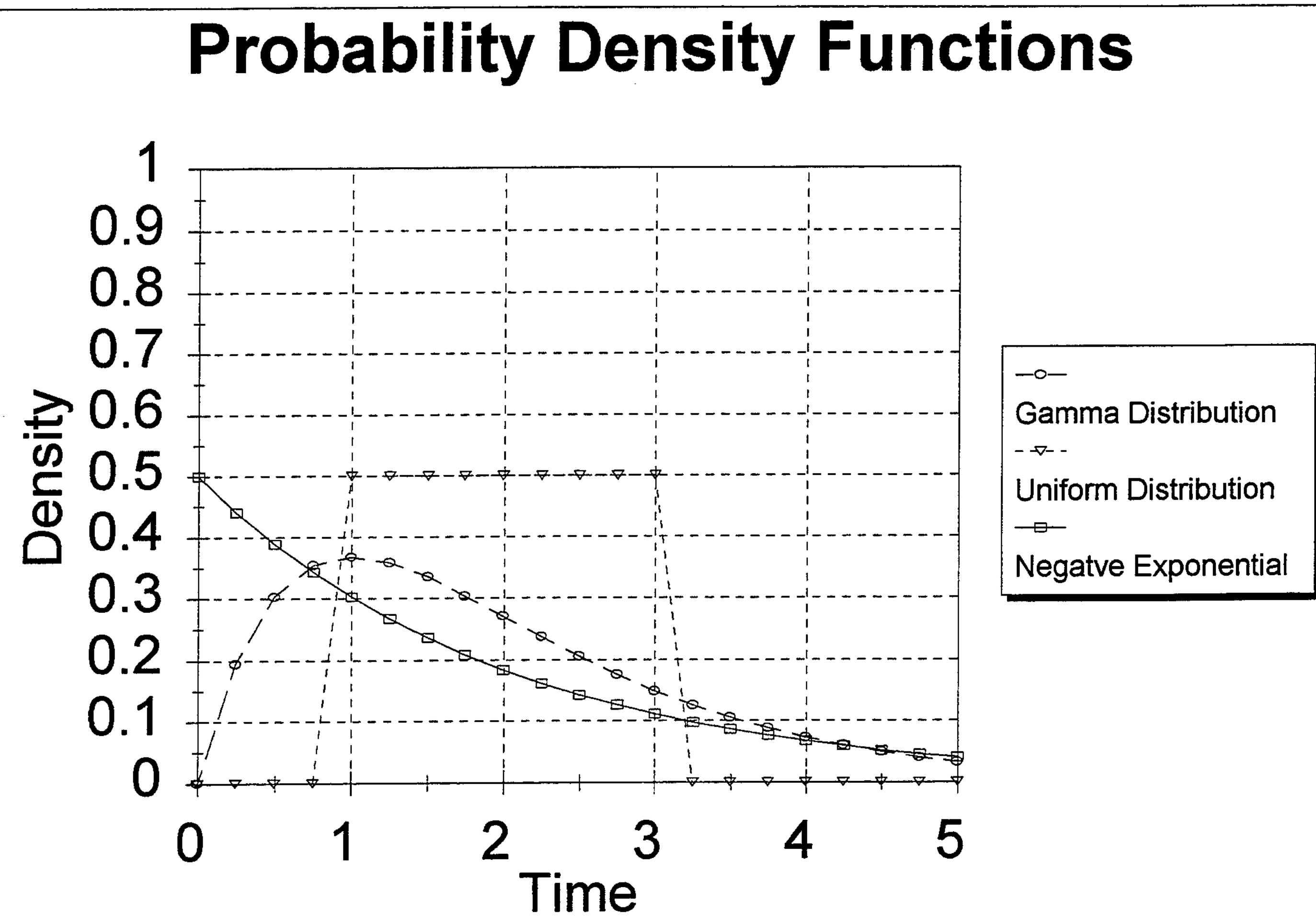


Figure XX.C.1. Probability Density Functions

These are shown in Figure XXI.C.2.

There are three features about this figure (i.e. about the data,) that we need to make. First, we note that for the uniform probability density function data, no losses occur until after t = 1. This is an artifact of the probability density function, since a firing time of less than 1 is not possible because of the lower limit on the probability density function.

Next, we note that the durations of the trajectories are different. This is an artifact of the way the algorithm was implemented into the simulation. Data sets were generated until limits on losses or time were reached. The calculations shown in this figure were then performed out to the least of the last event times of the data sets for that probability density function.

Lastly, the longest time points on the figure show greater variance than most of the points. This is also a result of the way the data set calculations are implemented (and ended) in the program.

Having said all of this, we may now come down to looking at the figure itself. Since all of these trajectories are based on probability density functions that have the same expectation value, we would expect all of them to have about the same rate (of decrease). Allowing for the shift implicit to the uniform probability density function, we must observe that this appears to be the case, although there is some spread in the data - the curves are clearly related but are also clearly different. Part of this difference must arise from experimental error- the limited sizes and number of the data sets, but part must arise from the higher order moment differences among the probability density functions. Taking these into account, we may still investigate these data sets.

Using the techniques embodied in equations (XXI.B-4)-(XXI.B-8) and equations (XXI.B-16)-(XXI.B-22), we may calculate attrition rates. Actually, we deviate from these to assume that Quadratic Lanchester attrition holds, so we calculate attrition rates using attrition differential equations,

$$\frac{dA}{dt} = -\alpha B , \qquad \text{(XXI.C-2)}$$

and

$$\frac{dB}{dt} = -\beta A , \qquad \text{(XXI.C-3)}$$

with two types of attrition rate coefficients,

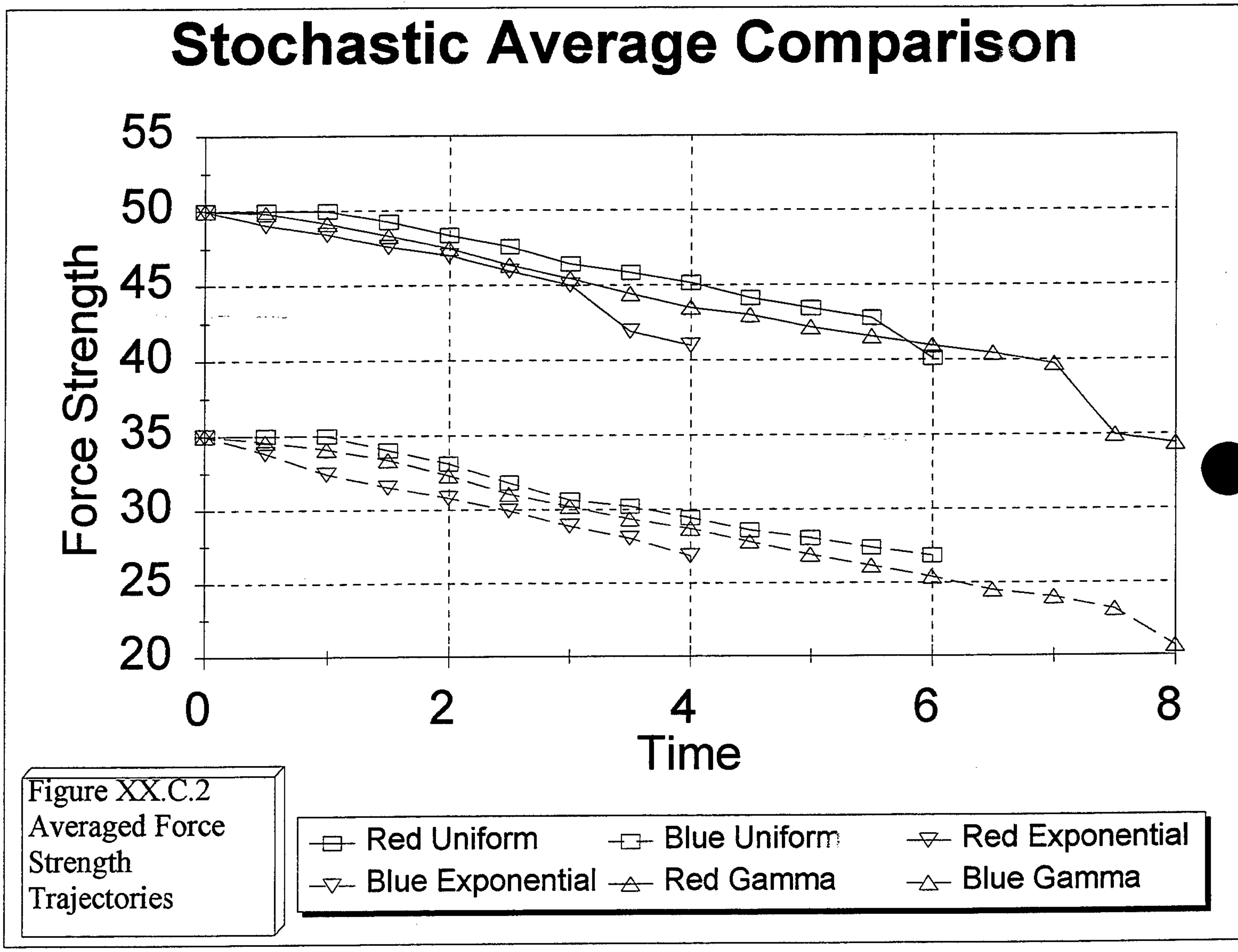


Figure XX.C.2
Averaged Force
Strength
Trajectories

$$\alpha_1 = \frac{1}{\langle \Delta t \rangle} , \qquad \text{(XXI.C-4)}$$

and

$$\alpha_1 = \langle \frac{1}{\Delta t} \rangle , \qquad \text{(XXI.C-5)}$$

where Δt are the inter-event times, and $\langle\rangle$ indicates that the Δt are averaged over all data sets for all events that occur in all data sets. The latter restriction is intended to prevent biasing the attrition rate coefficients due to reduced event occurrence.

Unfortunately, the variance in the calculations of α_2 (and β_2) are so great as to be useless, so we will not include these here. We merely note that this does not bode well for this type of attrition rate coefficient.

In addition, we calculated theoretical values for these attrition rate coefficients using the probability density functions and the $p_{k|s}$. We defer the theory behind these calculations till the next chapter, with apology to the student. This is unavoidable since the purpose of this chapter is to provide a theoretical and practical basis for rate theories of attrition.

These attrition rates are summarized in Table XXI.C.1 where the T subscripts indicate theoretical values..

Table XXI.C.1 Attrition Rate Coefficients

Prob. Dist. Fnc.	α_1	α_{1T}	α_{2T}	β_1	β_{1T}	β_{2T}
Uniform	0.058	0.050	0.055	0.036	0.035	0.038
NED	0.053	0.050	0.075	0.040	0.035	0.053
Gamma	0.060	0.050	0.100	0.036	0.035	0.070

The type 2 rate coefficients α_{2T} and β_{2T} were calculated approximately using a second order expectation value expansion.[1] Clearly, the NED rate

$$\langle \frac{1}{t} \rangle = \alpha \int_0^\infty \frac{dt}{t} e^{-\psi t} , \qquad \text{(XXI.C-6)}$$

is not trivial. We approximate the attrition rate coefficients by

$$\frac{1}{t} \simeq \frac{1}{\langle t \rangle} + \frac{t - \langle t \rangle}{\langle t \rangle^2} + \frac{1}{2} \frac{(t - \langle t \rangle)^2}{\langle t \rangle^3} + \text{HOT} , \qquad \text{(XXI.C-7)}$$

which gives

$$\langle \frac{1}{t} \rangle \simeq \frac{1}{\langle t \rangle} + \frac{1}{2} \frac{\sigma_t^2}{\langle t \rangle^3} . \qquad \text{(XXI.C-8)}$$

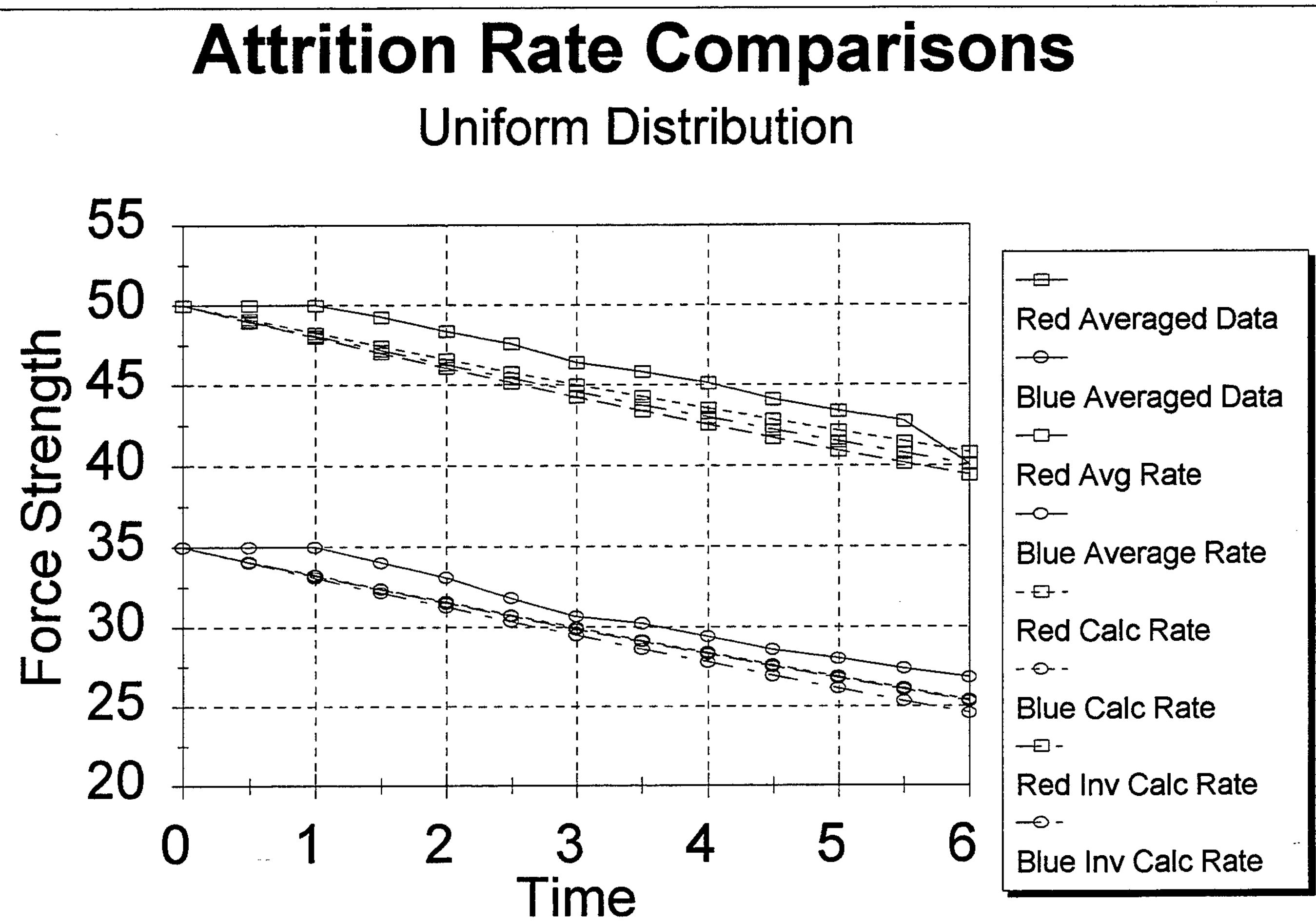
Since the value of the expectation will only increase with more terms, this value of α_{2T} is a minimum estimation. Even in this case, the values of these attrition rate coefficients are considerably larger than those calculated from the data.

Further, we see considerable consistency among the calculated attrition rates and what looks like reasonable agreement with the theoretical calculations. To examine this, we compare force strengths trajectory calculations using these attrition rate coefficients to the average force strength trajectories shown in the figure. We show these comparisons in Figures XXI.C.3. - XXI.C.5. In all cases, we find the calculations using α_{2T} and β_{2T} to have poor agreement, and the calculations using α_{1T} and β_{1T} to agree better with the data than those based on the data averaged coefficients. This indicates that the errors in the data may not be present in the theory.

We may make another comparison. Using the event times, we may compute average force strength trajectories using the averaged event times. Then , we may compare these to the calculations using the coefficients. These are shown in Figures XXI.C.6. - XXI.C.11.

If we examine the data for the uniform distribution, we find relatively poor agreement, primarily due to the shift from the distribution. For the NED, we find good agreement for the type 1 coefficient calculations, but poor agreement for the type 2 coefficient calculations. This situation is repeated for the gamma distribution case.

From this, we may conclude that the data support the type 1 rate theory, but not the type 2 rate theory. There is, however, one more comparison that we may make. Using all of the data sets, we may calculate average force strength trajectories of the form of equation (XXI.B-5) and compare these with calculations. While



Attrition Rate Comparisons
Uniform Distribution
Force Strength
55
50
45
40
35
30
25
20
0
1
2
3
4
5
6
Time
Red Averaged Data
Blue Averaged Data
Red Avg Rate
Blue Average Rate
Red Calc Rate
Blue Calc Rate
Red Inv Calc Rate
Blue Inv Calc Rate

Figure XX.C.3

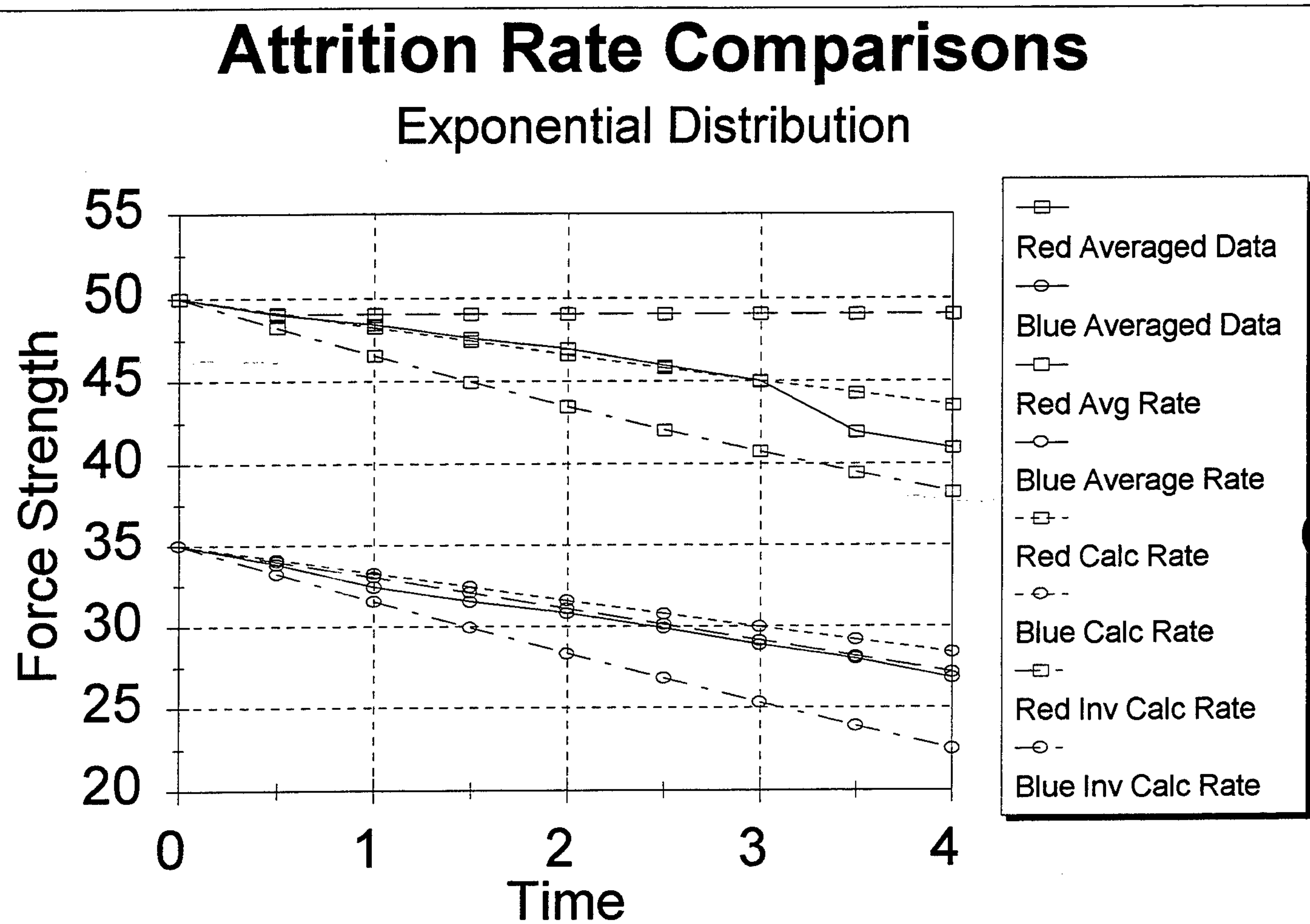
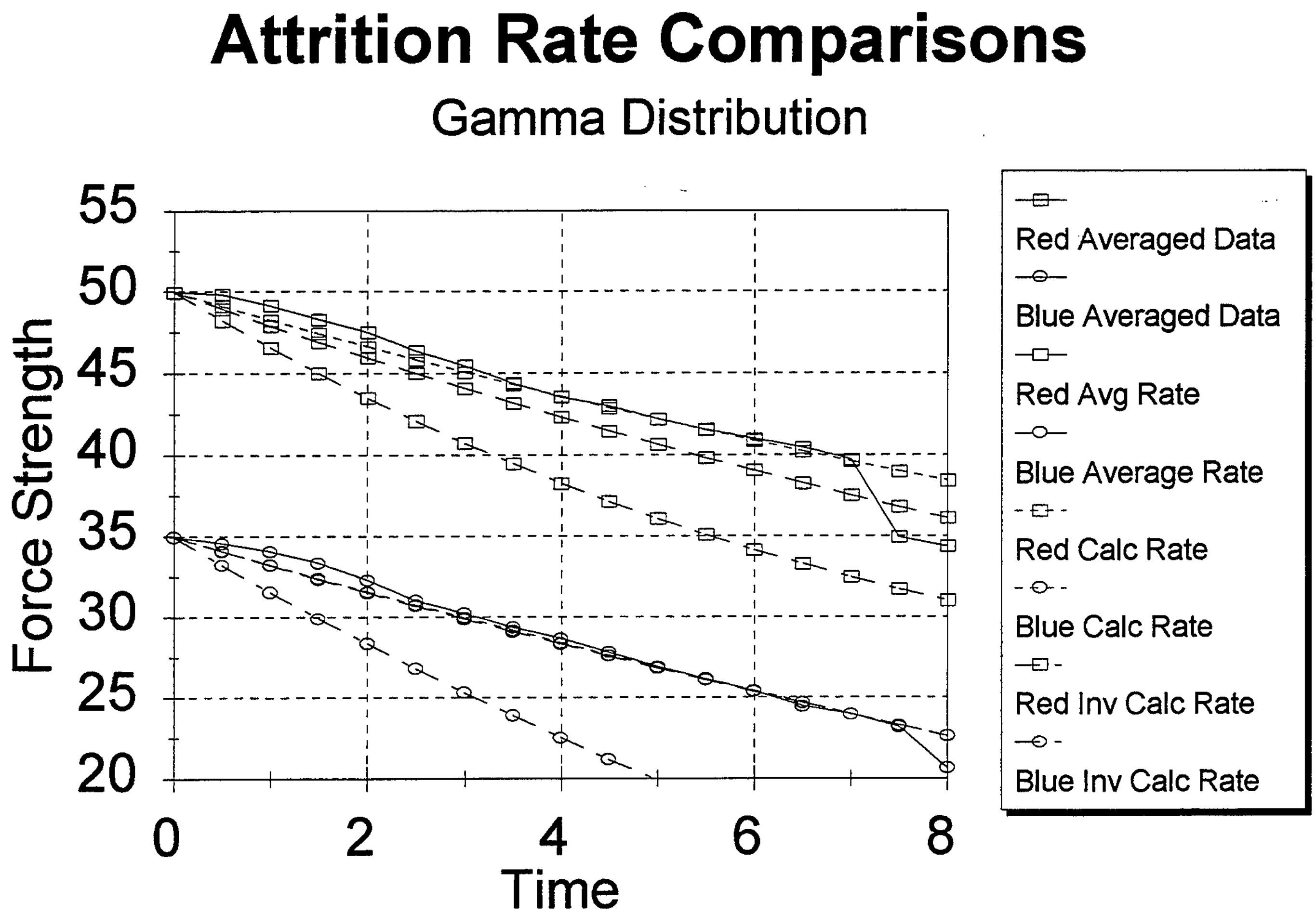
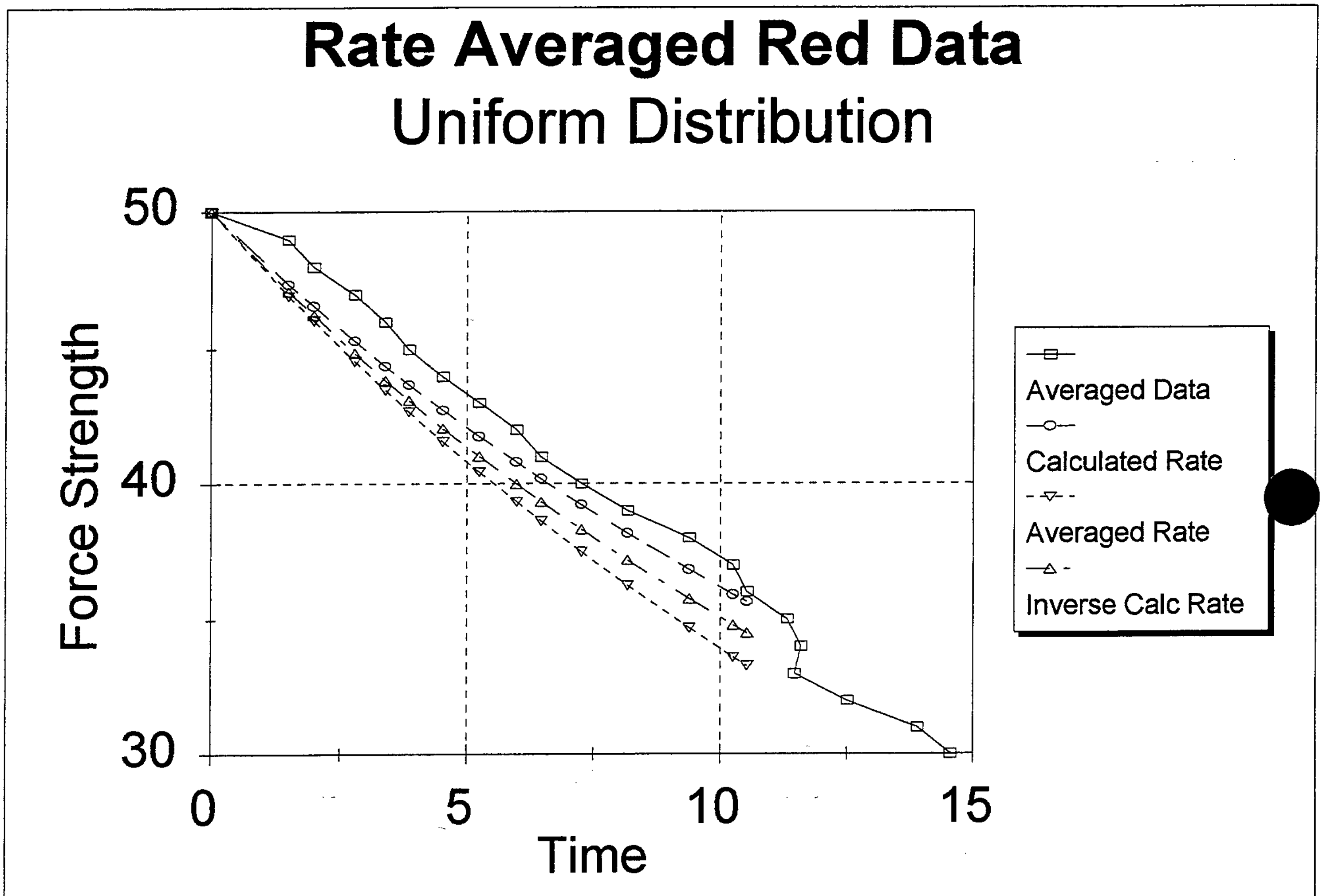

Attrition Rate Comparisons
Exponential Distribution
Force Strength
55
50
45
40
35
30
25
20
0
1
2
3
4
Time
Red Averaged Data
Blue Averaged Data
Red Avg Rate
Blue Average Rate
Red Calc Rate
Blue Calc Rate
Red Inv Calc Rate
Blue Inv Calc Rate

Figure XX.C.4



Attrition Rate Comparisons
Gamma Distribution
Force Strength
55
50
45
40
35
30
25
20
0
2
4
6
8
Time
Red Averaged Data
Blue Averaged Data
Red Avg Rate
Blue Average Rate
Red Calc Rate
Blue Calc Rate
Red Inv Calc Rate
Blue Inv Calc Rate

Figure XX.C.5



Rate Averaged Red Data
Uniform Distribution
50
40
30
Force Strength
0
5
10
15
Time
Averaged Data
Calculated Rate
Averaged Rate
Inverse Calc Rate

Figure XX.C.6.

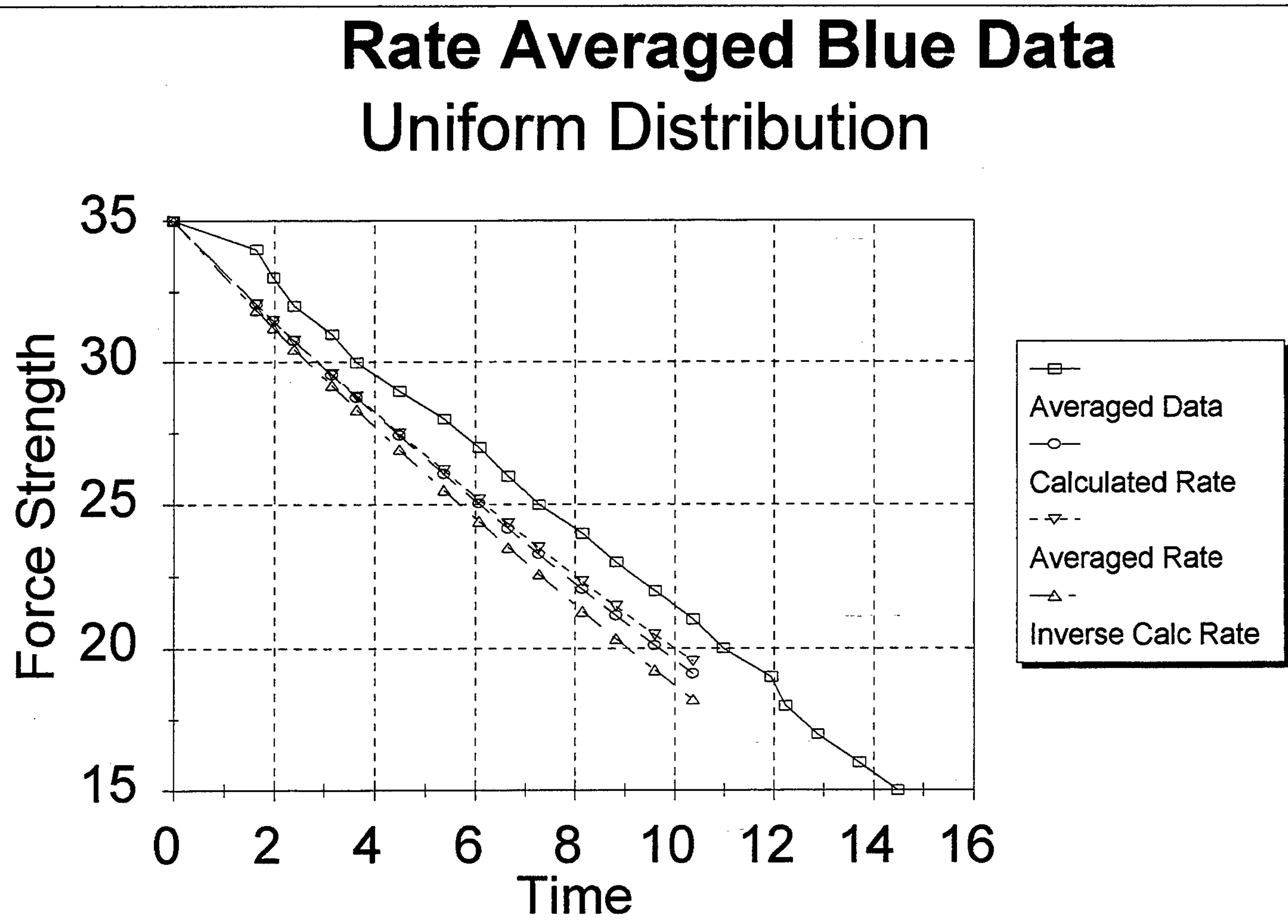

Rate Averaged Blue Data
Uniform Distribution
Force Strength
35
30
25
20
15
0
2
4
6
8
10
12
14
16
Time
Averaged Data
Calculated Rate
Averaged Rate
Inverse Calc Rate

Figure XX.C.7

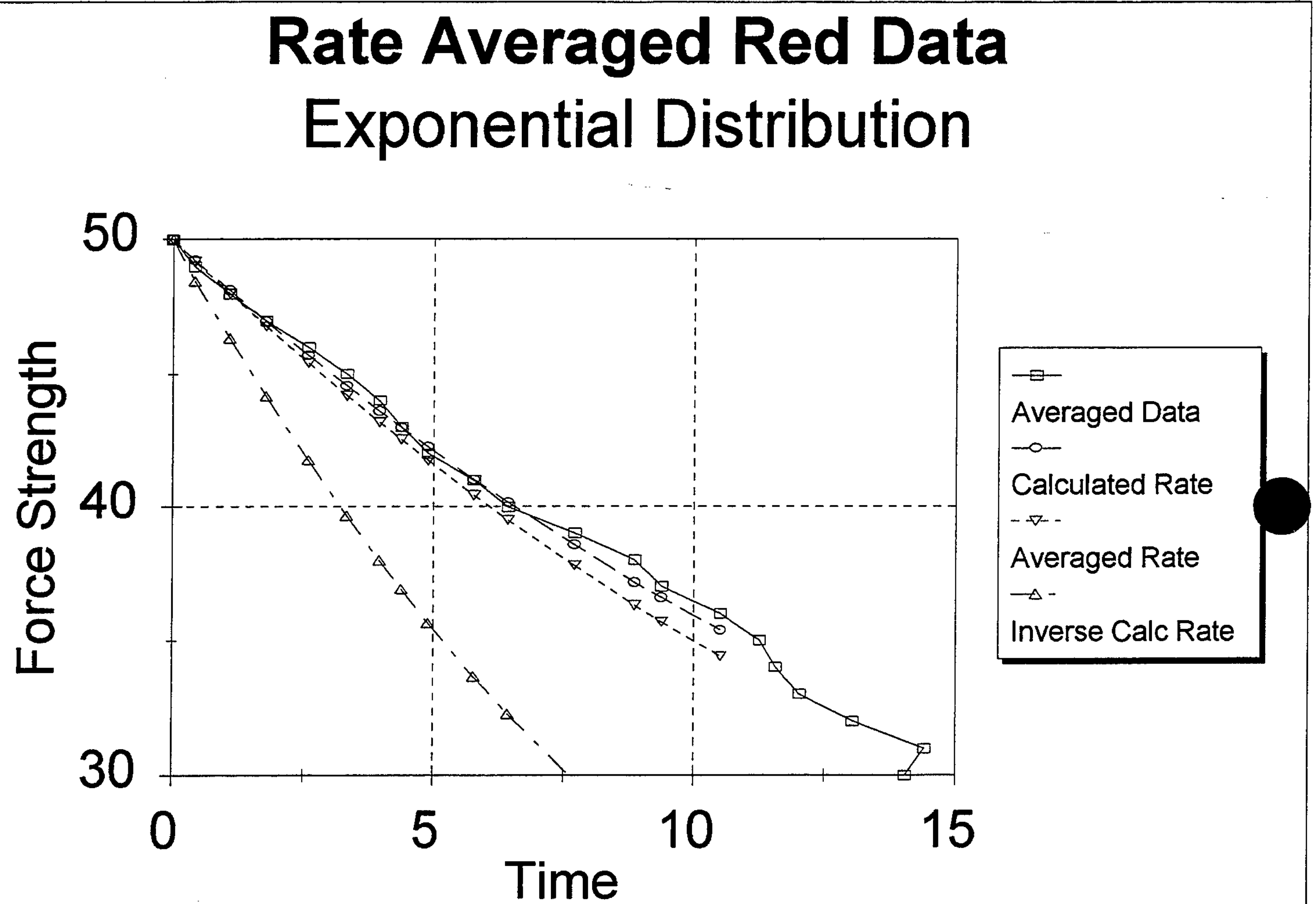
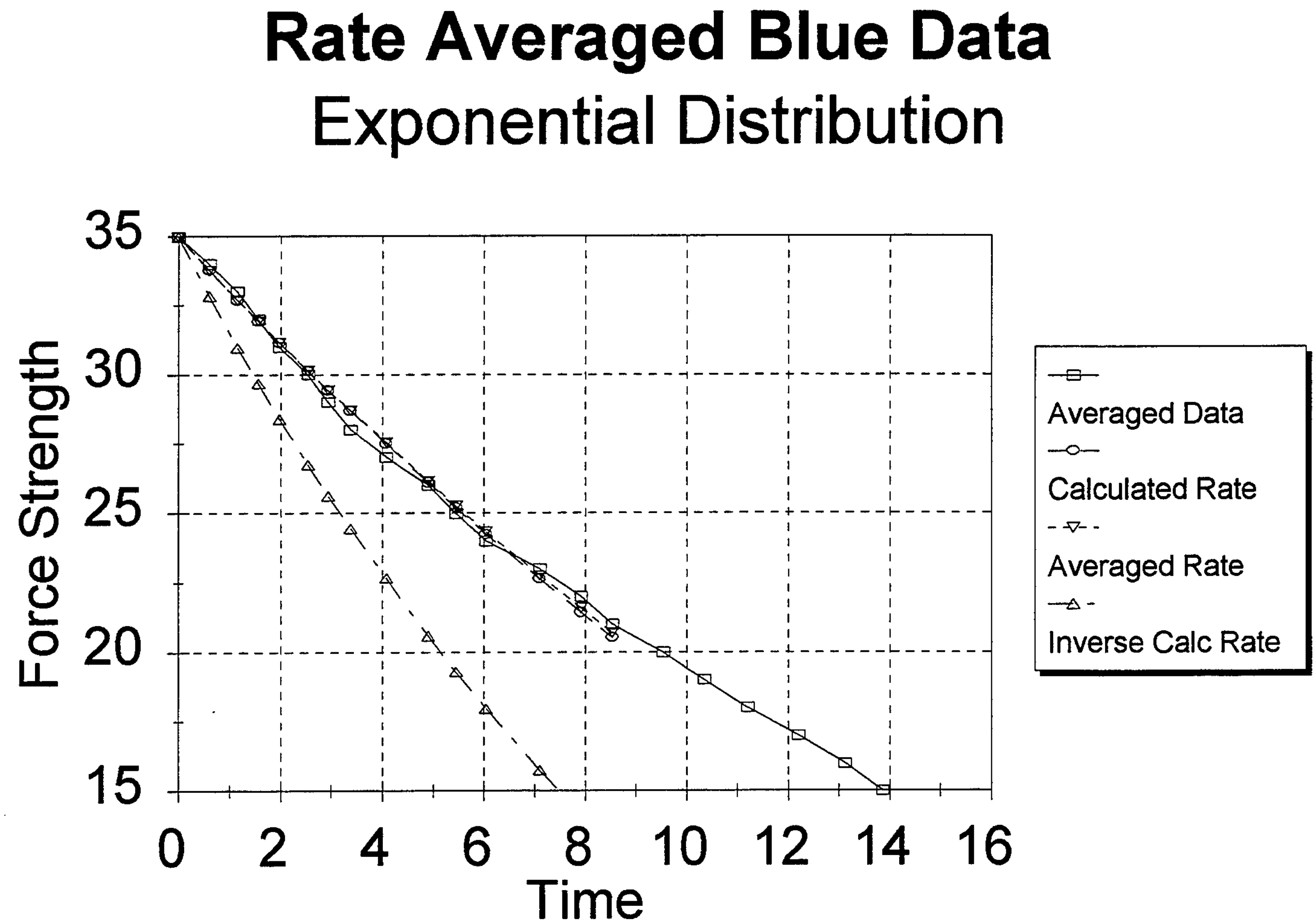

Rate Averaged Red Data
Exponential Distribution
50
40
30
Force Strength
0
5
10
15
Time
Averaged Data
Calculated Rate
Averaged Rate
Inverse Calc Rate

Figure XX.C.8.



Rate Averaged Blue Data
Exponential Distribution
Force Strength
35
30
25
20
15
0
2
4
6
8
10
12
14
16
Time
Averaged Data
Calculated Rate
Averaged Rate
Inverse Calc Rate

Figure XX.C.9.

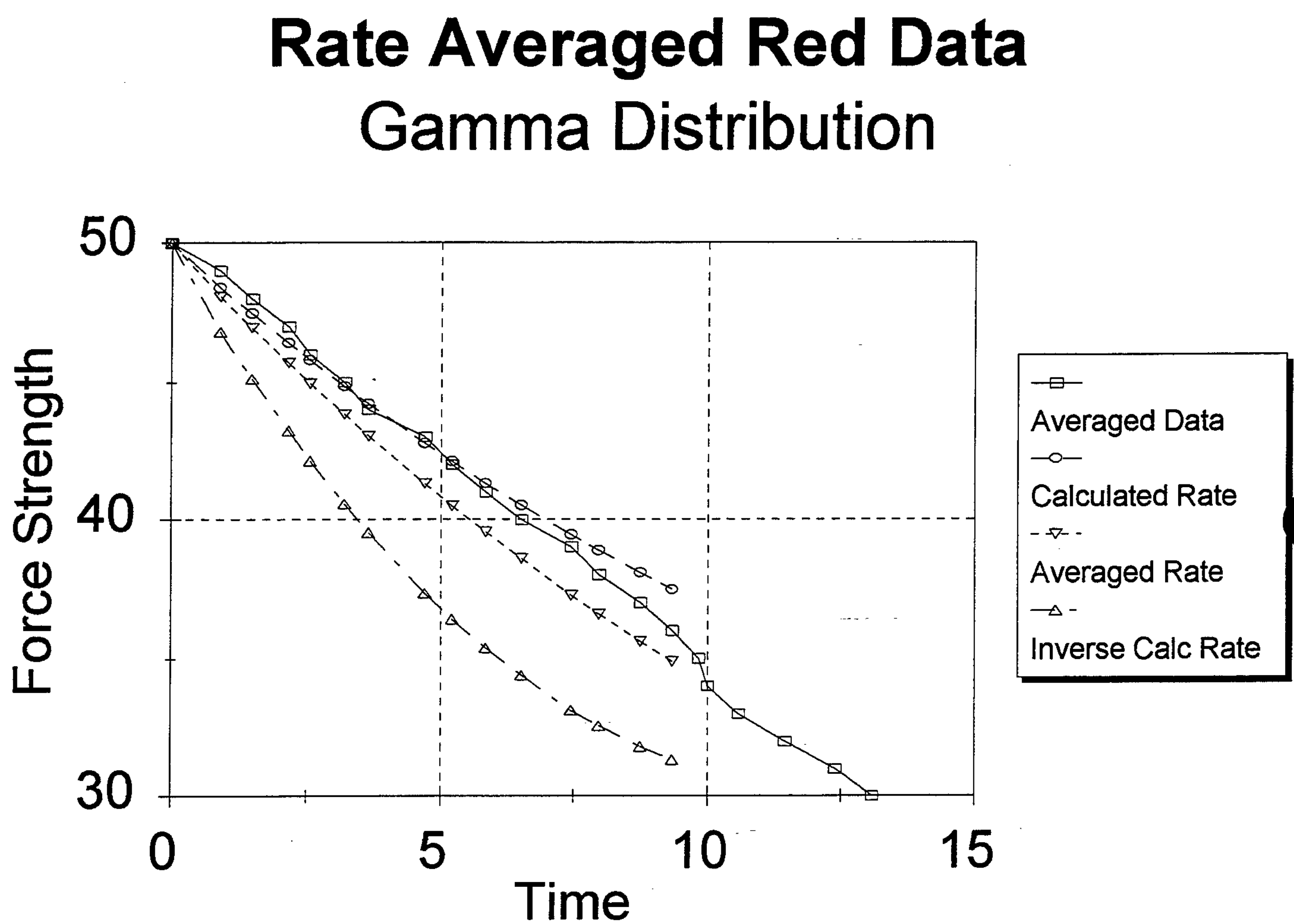


Figure XX.C.10.

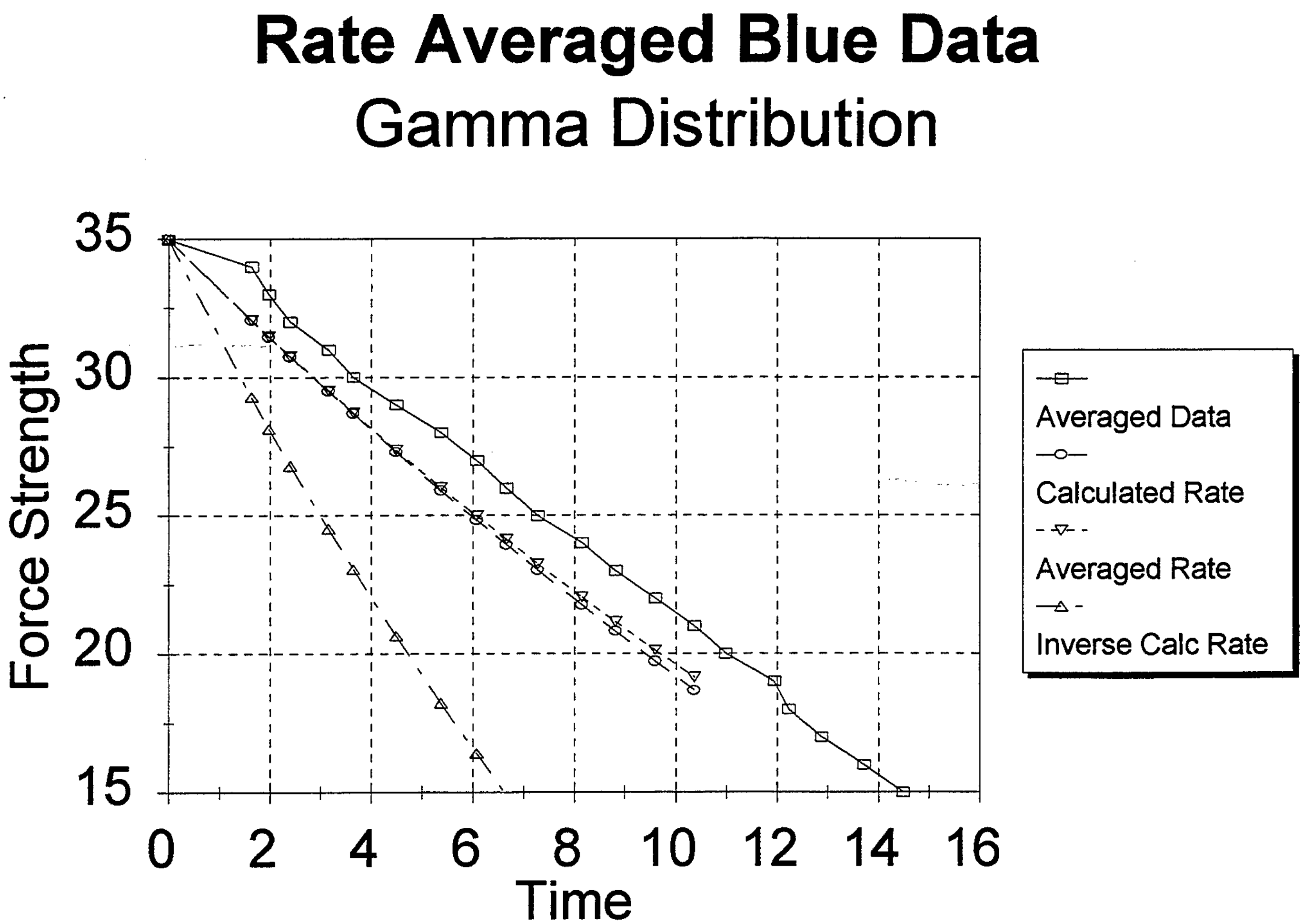


Figure XX.C.11.

the basic calculations are straight forward enough in our simulation, the actual manipulation of the data in a spreadsheet is manually interim. (At least for me, there is probably a smart way to do this, but I ended up doing it manually.) Further, even if we truncate the results at a short time (as we did in the Stochastic Average Comparison,) there is a lot of data. Besides being cumbersome to deal with, it is very slow to plot. Accordingly, we present only one set of these data, for the exponential distribution. We select this distribution because the exponential distribution is the one most often used (and criticized). We loose little from this selection, since our purpose is illustration. The data, and a pair of approximate calculations are shown in Figure XXI.C.12. The attrition rate coefficients used for the calculations are $\alpha = 0.056$ and $\beta = 0.045$. While the value of α is consistent with what we have already seen, the value of β is quite large.

D. Conclusion

In this chapter, we have laid the basis for considering the theory of rate processes, using the (ersatz) experimental data approved to compare a rate process with the stochastic general renewal theory approach. Even though our data are simulated, they are based on a model that is common to both theories. It is my intent (and my hope,) that this exposition will serve two purposes for the student.

First, it should provide some insight into the fundamental differences between the two theoretical approaches to modeling of combat and thereby an appreciation of the fact that these are subtle mathematical differences between the mathematical results of the two theories. In addition, the comparisons shown here provides a graphical demonstration of the differences between specific data and theoretical calculations. It indicates that even without the insertion of other stochastic combat complexities, there is considerable difference between what we might expect to observe in combat and what we would calculate. This gives us a graphical picture of what our expectations should be from Lanchester theory.

At this point, we may now turn our attention to a very different part of Lanchester theory. To this point, we have been primarily concerned with the basics of Lanchester attrition differential equations with some side trips thrown in to compare alternatives to and embellishments of the basics. We shall return to these considerations later, but first we will explore a crucial enabling conjugate to basic Lanchester attrition theory. This body of conjugate theory, most commonly typified by what is called Bonder-Farrell Theory, after its developers. This theory provides the basis for the calculation of attrition rates (and their coefficients/functions). This theory is the connection between the mechanics of combat, described by the attrition differential equations, and the physical and psychological mechanics of the combatants and their combat systems.

This connection is the empowerment of the Lanchester rate theory of combat

attrition. This is the promise of the theory; its ability to incorporate very complex processes, both physical and psychological, in an orderly mathematical manner; that is its inherent value as a descriptive tool. The very mathematical complexity of the stochastic general renewal theory, especially for large force sizes, precludes its use with these complex processes. Similarly, while computer technology has finally reached a point where simulations incorporating these complexities can be executed in reasonable time, replication is still required. Thus, the combination of Lanchester attrition differential equations (rate equation,) and conjugate attrition rate theory provide a general tool for the rapid consideration of a wide variety of combat situations and processes.

There are prices to be paid for this tool. Obviously, we will incur a loss of accuracy (or faithfulness) as we have indicated in this chapter. We must pay the price of understanding the models that contribute to the calculation of the attrition rates.

Finally, we shall find that the attrition rates we calculate with conjugate theory are not generally as simple as the ones we have considered in basic Lanchester theory. The resulting attrition differential equations are more complicated than before, and often (even generally,) do not have obvious analytical solutions. Accordingly, we must resort to numerical or simulation techniques to compute solutions and thereby insight is more difficult to draw from these solutions. Happily, however, these differential equations are usually very well behaved so that simple numerical techniques in spreadsheet simulations are still eminently feasible and practicable. The methods we have developed thus far will often serve us in good stead.

References

1. Sveshnikov, A. A., ed., **Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions,** Dover Publications, INC., New York, 1968.

APPENDICES

APPENDIX A. Useful Integrals

Several integrals are useful in solving some of the problems described here. As a convenience, we reproduce these integrals in this appendix as an aid to the reader. The source of the integrals is indicated.

Several indefinite integrals are of use in integrating the attrition differential equations.

$$\int \frac{dx}{x(ax-b)} = -\frac{2}{b}\coth^{-1}\left(\frac{2ax}{b} - 1\right) = \frac{1}{b}\ln\left(\frac{x}{ax+b}\right) \tag{A-1}$$

[Petit Bois, p. 2]

$$\int \frac{dx}{\sqrt{ax^2 + b}} = \frac{1}{\sqrt{a}}\sinh^{-1}\left(x\sqrt{\frac{a}{b}}\right) \tag{A-2}$$

[Petit Bois, p. 45]

$$\int \frac{dx}{b - ax^2} = \frac{1}{\sqrt{ab}}\tanh^{-1}\left(x\sqrt{\frac{a}{b}}\right) \tag{A-3}$$

[Petit Bois, p. 2]

$$\int \frac{dx}{x\sqrt{ax+b}} = -\frac{2}{b}\coth^{-1}\left(\sqrt{\frac{ax}{b}+1}\right) = \frac{1}{\sqrt{b}}\ln\left(\frac{\sqrt{ax+b} - \sqrt{b}}{\sqrt{ax+b} + \sqrt{b}}\right) \tag{A-4}$$

[Petit Bois, p. 2]

$$\int \frac{dx}{ax^2 - bx} = -\ln\left(\frac{ax-b}{x}\right) \tag{A-5}$$

APPENDIX A. Useful Integrals

$$\int d\coth^{-1}(x) = \int \frac{dx}{1-x^2}, \; x^2 < 1 \tag{A-6}$$

$$\int d\tanh^{-1}(x) = \int \frac{dx}{1-x^2}, \; x^2 > 1 \tag{A-7}$$

$$\int d\,\mathrm{sech}^{-1}(x) = \mp \int \frac{dx}{x\sqrt{1-x^2}} \tag{A-8}$$

$$\int d\,\mathrm{csch}^{-1}(x) = \mp \int \frac{dx}{x\sqrt{1+x^2}} \tag{A-9}$$

$$\int_0^\infty e^{-x^\mu}\, dx = \frac{1}{\mu}\Gamma(\mu). \tag{A-10}$$

[Gradshteyn and Ryzhik, 3.326]

$$\int_0^x \int_0^{x_n} \cdots \int_0^{x_3} \int_0^{x_2} f(x_1)\, dx_1\, dx_2 \ldots dx_{n-1}\, dx_n = \frac{1}{(n-1)!} \int_0^x (x-\zeta)\, f(\zeta)\, d\zeta\,. \tag{A-11}$$

[Hildebrand, pp. 224-225]

APPENDIX B. Useful Formulae

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} \ldots, \; x < 1 \qquad \text{(B-1)}$$

$$\tanh(x+y) = \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(x)} \qquad \text{(B-2)}$$

$$\coth(x+y) = \frac{1 + \coth(x)\coth(y)}{\coth(x) + \coth(y)} \qquad \text{(B-3)}$$

$$\tanh^{-1}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right), \; x^2 < 1 \qquad \text{(B-4)}$$

$$\coth^{-1}(x) = \frac{1}{2}\ln\left(\frac{x+1}{x-1}\right), \; x^2 > 1 \qquad \text{(B-5)}$$

Stirling's Approximation

$$\ln(N!) \simeq N\ln(N) - N \qquad \text{(B-6)}$$

APPENDIX C. Alternate Forms of the Attrition Solution

The Quadratic Lanchester Attrition Solutions have the form:

$$A(t) = A_0 \cosh(\gamma \Delta t) - \delta B_0 \sinh(\gamma \Delta t) \quad \text{(C-1)}$$

and

$$B(t) = B_0 \cosh(\gamma \Delta t) - \frac{A_0}{\delta} \sinh(\gamma \Delta t) \quad \text{(C-2)}$$

where:

$$\gamma = \sqrt{\alpha \beta}\ , \quad \text{(C-3)}$$

and

$$\delta = \sqrt{\frac{\alpha}{\beta}}\ . \quad \text{(C-4)}$$

It is convenient to rewrite these in exponential form as:

$$\sqrt{\beta} A(t) = \frac{\left(\sqrt{\beta} A_0 - \sqrt{\alpha} B_0\right) e^{\gamma \Delta t} + \left(\sqrt{\beta} A_0 + \sqrt{\alpha} B_0\right) e^{-\gamma \Delta t}}{2}\ , \quad \text{(C-5)}$$

and

$$\sqrt{\alpha} B(t) = \frac{\left(\sqrt{\alpha} B_0 - \sqrt{\beta} A_0\right) e^{\gamma \Delta t} + \left(\sqrt{\alpha} B_0 + \sqrt{\beta} A_0\right) e^{-\gamma \Delta t}}{2}\ , \quad \text{(C-6)}$$

These two equations amply demonstrate the dependence of the solutions on the draw situations (combat to a conclusion.)

APPENDIX D. Historical Databases

Data Set	Table No.	Number of Battles
Nominal Length Battles	D.1	108
Civil War Battles	D.2	49
Osipov's Battles	D.3	45
Short Battles	D.4	72
World War I Battles	D.5	12

In this appendix we explicitly present the databases introduced in Chapter 9. These databases are:

Appendix D

Battle	Date	Blue	Blue Start	Blue Finish	Blue Loss	Red	Red Start	Red Finish	Red Loss	Duration
Marathon	-490	Greeks	11000	10808	192	Persians	20000	13600	6400	1
Hydaspes	-327	Macedonians	135000	134000	1000	Indians	30000	18000	12000	1
Arbela	-311	Persians	245000	155000	90000	Macedonians	47000	46500	500	1
Magnesia	-190	Syrians	80000	30000	50000	Romans	40000	39700	300	1
Mursa	351	Rebels	100000	76000	24000	Romans	80000	50000	30000	1
Navarette	1367	British	20000	19900	100	French/Spanish	40000	33000	7000	1
Marignano	1575	French	50000	44000	6000	Swiss	40000	29200	10800	2
Breitenfeld I	1631	Swedes	40000	37300	2700	Imperials	44000	36000	8000	1
Wittstock	1636	Swedes	22000	17000	5000	Austrians/Spani	30000	18000	12000	1
Zurakow	1676	Poles	100000	97000	3000	Turks	200000	120000	80000	185
Ramilles	1706	French	50000	35000	15000	Allies	50000	47000	3000	1
Oudenarde	1708	French	100000	94000	6000	Allies	78000	75000	3000	1
Mollwitz	1741	Prussians	30000	27500	2500	Austrians	30000	25000	5000	1
Sugar-Loaf Rock	1753	French/Indians	34000	33900	100	British/Indians	6000	5960	40	1
Lake George	1755	French/Indians	1500	1100	400	American Coloni	2500	2098	402	1
Breslau	1757	Austrians	90000	84000	6000	Prusians	25000	20000	5000	1
Plassey	1757	French/Indians	50000	49500	500	British/Indians	3000	2928	72	0.2
Ticonderoga I	1758	French	3600	3223	377	British/America	15000	13056	1944	1
Madras II	1758	French/Indians	6000	4900	1100	British/Indians	4000	2659	1341	60
Zerndorf	1758	Russians	42000	21000	21000	Prussians	36500	22700	13800	1
Plains of Abrah	1759	British	4000	3336	664	French	4000	2500	1500	1
Trincomalee II	1767	British/Indians	12000	8000	4000	Mysore Forces	60000	59840	160	1
Quebec II	1775	British	1800	1782	18	Americans	600	500	100	1
Brooklyn	1776	Americans	11000	9000	2000	British	30000	29690	310	1
Moore's Creek B	1776	Americans	1100	1098	2	British	1800	1770	30	1
Brandywine	1777	British	18000	17410	590	Americans	8000	7100	900	1
Saratoga	1777	British	6000	5400	600	Americans	7200	7000	200	1
Newport	1778	Americans	10000	9689	311	British	3000	2740	260	23
Stony Point	1779	British	700	567	133	Americans	1350	1255	95	2
Panani	1780	British	2500	2413	87	Mysore Forces	18300	17200	1100	1
Hobkirk's Hill	1781	Americans	1551	1281	270	British	900	642	258	1
Sholinghar	1781	British	10000	9900	100	Mysore Forces	80000	75000	5000	1
Martinesi	1789	Turks	80000	70000	10000	Austrians/Russi	27000	26383	617	1
Savannah II	1799	Americans/Frenc	5050	4212	838	British	3200	3045	155	1
Engen	1800	French	75000	73000	2000	Austrians	110000	108000	2000	1
Stekach II	1800	French	50000	48000	2000	Austrians	60000	58000	2000	21
Pultusk II	1806	Russians	37000	34000	3000	French	20000	15800	4200	1
Rio Seco	1808	French	14000	13630	370	Spanish	26000	20000	6000	1
Talavera	1809	British/Spanish	40000	34600	5400	French	50000	42700	7300	1
Aspern-Essling	1809	French	55000	35000	20000	Austrians	90000	67000	23000	2

Table D.1. Nominal Length Battles

Appendix D

Battle	Date	Blue	Blue Start	Blue Finish	Blue Loss	Red	Red Start	Red Finish	Red Loss	Duration
Eckmuhl	1809	French	90000	85000	5000	Austrians	76000	69000	7000	1
Raab	1809	French	44000	41500	2500	Austrians	40000	37000	3000	1
Wagram	1809	French	190000	156000	34000	Austrians	139000	96000	43000	3
Medelin	1809	French	17500	16500	1000	Spansih	30000	12000	18000	1
Busaco	1810	British	25000	23500	1500	French	40000	35500	4500	1
Albuhera	1811	French	33000	25000	8000	Allies	7000	1800	5200	1
Mohilev	1812	French	28000	27000	1000	Russians	60000	56000	4000	1
Smolensk II	1812	French	50000	41000	9000	Russians	60000	50000	10000	1
Malo Jaroslavet	1812	Russians	24000	18000	6000	French	15000	10000	5000	1
Maya	1813	British	6000	4600	1400	French	26000	24500	1500	1
Sauroren	1813	French	25000	22000	3000	British	12000	9400	2600	1
Bautzen	1813	French	150000	148700	1300	Prussians/Russi	100000	85000	15000	3
Baltimore	1814	British	3270	2924	346	Americans	17000	16890	110	1
Lundy's Lane	1814	British	3000	2150	850	Americans	5000	4150	850	0.2
Toulose II	1814	British/Spanish	25000	20400	4600	French	30000	27000	3000	1
La Rotheria	1814	French	32000	27000	5000	Allies	100000	92000	8000	1
Arcis sur Aube	1814	French	23000	21300	1700	Austrians	60000	57500	2500	1
Craonne	1814	Prussians	90000	85000	5000	French	37000	31600	5400	1
Quatre Bras	1815	French	25000	20700	4300	British/Dutch	36000	31300	4700	1
Mahidput	1817	Indians	35000	32000	3000	British	5500	4722	778	1
Alamo	1836	Mexicans	2500	900	1600	Texans	185	0	185	1
Maharajapore	1843	Indians	18000	15000	3000	British	40000	39213	787	1
Resaca de la Pa	1846	AMericans	1700	1578	122	Mexicans	5100	4483	617	1
Buena Vista	1847	Americans	4500	3754	746	Mexicans	18000	16500	1500	1
Vera Cruz	1847	Mexicans	5000	4820	180	Americans	13000	12918	82	6
Cerro-Gordo	1847	United States	8500	8100	400	Mexicans	12000	11300	700	1
Alma River	1854	Russians	40000	34100	5900	British/French	26000	22000	4000	1
Inkerman	1854	Russians	50000	38000	12000	British/French	8000	4500	3500	1
Bull Run I	1861	Union	40000	38508	1492	Confederates	30000	28018	1982	1
Wilson's Creek	1861	Union	5600	4364	1236	Confederates	11800	10705	1095	1
Richmond I	1862	Americans	8000	6954	1046	Confederates	6000	5540	460	1
Pea Ridge	1862	Confederates	16000	15200	800	Union	16000	14616	1384	2
Prarie Grove	1862	Confederates	11000	10000	1000	Union	10000	8700	1300	1
Seven Days	1862	Confederates	100000	80000	20000	Union	95000	79000	16000	7
Shiloh	1862	Confederates	43000	32303	10697	Union	42000	28913	13087	2
Fort Donelson	1862	Union	25000	22168	2832	Confederates	12000	10000	2000	1
Fredericksberg	1862	Union	150000	136229	13771	Confederates	80000	78200	1800	1
Front Royal	1862	Union	1063	159	904	Confederates	16000	15950	50	1
Secessionville	1862	Union	6000	5400	600	Confederates	2000	1800	200	1
Nashville	1863	Confederates	31000	29500	1500	Union	41000	37939	3061	1

Table [illegible]1. Nominal Length Battles

Appendix D

Battle	Date	Blue	Blue Start	Blue Finish	Blue Loss	Red	Red Start	Red Finish	Red Loss	Duration
Chancellorsvill	1863	Union	120000	102000	18000	Confederates	53000	43000	10000	3
Chattanooga	1863	Union	80000	74527	5473	Confederates	64000	61479	2521	4
Stones River	1863	Union	45000	32094	12906	Confederates	35000	23261	11739	4
Brices Cross Ro	1864	Confederates	3500	3008	492	Union	8000	7283	717	1
Monocacy River	1864	Confederates	14000	13300	700	Union	6000	4120	1880	1
Sabine Cross Ro	1864	Confederates	8300	5800	2500	Union	12500	9000	3500	2
Cold Harbor	1864	Union	108000	101000	7000	Confederates	59000	57500	1500	0.04
Mobile Bay	1864	Union	5500	5181	319	Confederates	470	158	312	3
Spotsylvania	1864	Union	101000	83600	17400	Confederates	56000	46400	9600	3
Wilderness	1864	Union	120000	102334	17666	Confederates	64000	56250	7750	5
Winchester III	1864	Union	38000	33117	4883	Confederates	12000	9897	2103	1
Custozza II	1866	Austrians	80000	75400	4600	Italians	140000	136168	3832	1
Mantana	1867	Italians	10000	8900	1100	Allies	5000	4818	182	1
Fort Kearney	1867	United States	32	25	7	Indians	1500	1300	200	1
Bel fort II	1871	French	150000	144000	6000	Germans	60000	58000	2000	3
Saint Quentin I	1871	French	40000	36500	3500	Germans	33000	30600	2400	1
Ulund	1879	Zulu	20000	18500	1500	British	5000	4907	93	1
Rorke's Drift	1879	Zulus	4000	3600	400	British	139	114	25	1
Bronkhurst Spru	1880	British	259	104	155	Americans	150	98	52	1
Son Tai	1883	Chinese	25000	24000	1000	French	7000	6590	410	3
Slivnitza	1885	Serbs	25000	22000	3000	Bulgars	15000	13000	2000	3
Omdurhman	1898	British	26000	25500	500	Mahdists	45000	30000	15000	1
Atbara	1898	British/Egyptia	14000	13430	570	Mahdists	18000	13000	5000	1
Mafeking	1900	Boers	5000	4000	1000	British	700	427	273	185
Mukden I	1905	Russians	300000	200000	100000	Japanese	300000	250000	50000	17
Tannenburg II	1914	Russians	300000	270000	30000	Germans	300000	287000	13000	5
Maypo	1918	Chileans	9000	8000	1000	Spanish	6000	5000	1000	1
Goose Green	1982	British	450	402	48	Argentines	1350	979	371	1

Table D.1. Nominal Length Battles

Appendix D

Battle	Union Start	Finish	Killed	Wounded	Confederate Start	Finish	Killed	Wounded	Duration
Bull Run	28452	26960	481	1011	32232	30263	387	1582	1
Wilson's Creek	5400	4456	223	721	11600	10443	257	900	1
Fort Donelson	27000	24392	500	2108	21000	19000	2000		4
Pea Ridge Arkan	11250	10067	203	980	14000	13400	600		1
Shiloh	62682	52520	1754	8408	40335	30600	1723	8012	1
Williamsburg	40768	38902	456	1410	31823	30253	1570		1
Fair Oaks	41797	37413	790	3594	41816	36087	980	4749	1
Mechanicsville	15631	15375	49	207	16356	14872	1484		1
Gaine's Mill	34214	30213	894	3107	57018	48267	8751		1
Peach Orchard,	83345	78376	724	4245	86748	78146	8602		2
Seven Day's Bat	91169	81373	1734	8062	95481	75742	3478	16261	6
Cedar Mountain	8030	6271	314	1445	16868	15530	231	1107	1
Manassas & Chan	75696	65600	1724	8372	48527	39419	1481	7627	6
Richmond KY	6500	5450	206	844	6850	6400	78	372	1
South Mountain	28480	26752	325	1403	17852	15967	325	1560	1
Antietam	75316	63659	2108	9549	51844	40120	2700	9024	1
Corinth	21147	18951	355	1841	22000	19530	473	1997	1
Perryville	36940	33244	845	2851	16000	12855	510	2635	1
Prairie Grove A	10000	9012	175	813	10000	9019	164	817	1
Fredericksburg	106007	95123	1284	9600	72497	67841	595	4061	1
Chickasaw Bayou	30720	29507	208	1005	13792	13595	63	134	2
Stone's River	41400	32180	1677	7543	34732	25493	1294	7945	1
Arkansas Post	28944	27912	134	898	4564	4455	28	81	1
Chancellorsvill	97382	86213	1575	9594	57352	46606	1665	9081	3
Champion Hill	29373	27119	410	1844	20000	17819	381	1800	1
Port Hudson Ass	13000	11162	293	1545	4192	3957	235		1
Port Hudson Ass	6000	4396	203	1401	3487	3440	22	25	1
Gettysburg	83289	65605	3155	14529	75054	52416	3903	18735	3
Fort Wagner Ass	5264	4138	246	880	1785	1616	36	133	1
Chickamauga	58222	46809	1657	9756	66326	49340	2312	14674	1
Chattanooga	56359	50884	753	4722	46165	43644	361	2160	2
Mine Run	69643	68371	173	1099	44426	43746	110	570	4
Olustee FLA	5115	3760	203	1152	5200	4266	93	841	1
Pleasant Hill	12647	11653	150	844	14300	13300	1000		1
Wilderness	101895	87612	2246	12037	61025	53275	7750		2
Drewry's Bluff	15800	13030	390	2380	18025	15729	355	1941	4
The Mine	20708	17844	2864		11466	10847	619		1
Weldon Railroad	20289	18986	198	1105	14787	13587	1200		3
Atlanta	110123	99595	10528		66089	56902	9187		28

Table 2. Civil War Battles

Appendix D

Battle	Union				Confederate				Duration
	Start	Finish	Killed	Wounded	Start	Finish	Killed	Wounded	
Kenesaw Mountai	16225	14226	1999		17733	17463	270		1
Tupelo	14000	13364	77	559	6600	5274	210	1116	2
Peach-Tree Cree	21655	20055	1600		18832	16332	2500		1
Atlanta	30477	28488	430	1559	36934	29934	7000		1
Atlanta	13226	12667	559		18450	14350	4100		1
Jonesborough GA	14170	13991	179	0	23811	22086	1725		1
Winchester	37771	33091	697	3983	17103	15000	276	1827	1
Cedar Creek	30829	26755	644	3430	18410	16550	320	1540	1
Franklin	27939	26717	189	1033	26897	21347	1750	3800	1
Bentonville	16127	15194	139	794	16895	15387	195	1313	1

Table D.2. Civil War Battles

Appendix D

Battle	Stronger Nation	Initial	Losses	Final	Weaker Nation	Initial	Losses	Final	Date
Kpaon	French	30	18	12	Russian	18	5	13	1814
Kul'm	Allies	46	9	37	French	35	10	25	1813
Auershtedt	Prussians	48	8	40	French	30	7	23	1806
Madzhenta	Austrian	58	10	48	French	54	5	49	1859
Aladzha	Russian	60	2	58	Turks	36	15	21	1877
Al'ma	Allies	62	3	59	Russian	34	6	28	1854
Chernaia	Allies	62	2	60	Russian	56	8	48	1854
Kustotsa	Austrian	70	8	62	Italians	51	8	43	1866
Dennevits	French	70	9	61	Allies	57	9	48	1813
Grokhoro	Russian	72	9	63	Poles	56	12	44	1831
Jena	French	74	4	70	Prussians	43	12	31	1806
Berezina	Russians	75	6	69	French	45	15	30	1812
Katsbach	Allies	75	3	72	French	65	12	53	1813
Aspern	Austrians	75	25	50	French	70	35	35	1809
Ganau	French	75	15	60	Allies	50	9	41	1813
Preisish	French	80	25	55	Russians	64	26	38	1807
Austerlitz	Allies	83	27	56	French	75	12	63	1805
Freiland	French	85	12	73	Russians	60	15	45	1807
Inkerman	Russian	90	12	78	Allies	63	3	60	1854
Laon	Allies	100	2	98	French	45	6	39	1814
Waterloo	Allies	100	22	78	French	72	32	40	1815
Vert	German	100	10	90	French	45	5	40	1870
Lun'i	French	120	11	109	Prussian	85	11	74	1815
Mars LaTour	French	125	16	109	Germans	65	16	49	1870
Borodino	French	130	35	95	Russians	103	40	63	1812
Liaoian	Russian	150	18	132	Japanese	120	24	96	1904
Liutsen	French	157	15	142	Allies	92	12	80	1813
Wagram	French	160	25	135	Austrians	124	25	99	1809
Drezden	Allies	160	20	140	French	125	15	110	1813
Bautsen	French	163	18	145	Allies	96	12	84	1813
Col'ferino	Austrian	170	20	150	French	150	18	132	1859
Mets	German	200	6	194	French	173	20	153	1870
Shabh	Russian	212	40	172	Japanese	157	20	137	1904
Gravelot	German	220	20	200	French	130	12	118	1870
Kenigrets	Prussians	222	10	212	Austrians	215	43	172	1866
Sedan	German	245	9	236	French	124	17	107	1870
Keiptsig	Allies	300	50	250	French	200	60	140	1813
Mukden	Russian	300	59	241	Japanese	280	70	210	1905

Table [illegible]3. Osipov's Battles

Date	Battle	Attacker	Attacker Initial	Final	Losses	Defender	Initial	Final	Losses	Duration Minimum	Maximum
1759	Quebec I	British	4500	3840	660	French	14000	12600	1400	0.25	0.25
1814	Chippewa River	Brown	1300	1025	275	Riall	1500	988	512	0.5	0.5
1815	New Orleans I	British	5300	3264	2036	Americans	4500	4479	21	0.5	0.5
1779	Kettle Creek	Pickens	300	268	32	Boyd	700	660	40	1	2
1898	La Guasimas	Wheeler	1800	1732	68	Spanish	1500	1465	35	1	2
1775	Bunker Hill	Howe	2500	1446	1054	Prescott	1600	1189	411	2	5
1862	Front Royal	Jackson	16000	15950	50	Kenly	1063	159	904	2	8
1864	Kenesaw Mt	Sherman	16000	14000	2000	McPherson	17000	16730	270	2	2
1864	Killdeer Mt	Sully	2200	2100	100	Sioux	5000	4985	15	2	8
1780	King's Mt	Shelby	900	580	320	Ferguson	1100	1010	90	2	8
1700	Narva	Swedes	8000	7000	1000	Russians	40000	30000	10000	2	3
1898	Omdurman	British	26000	25500	500	Mahdists	45000	30000	15000	2	4
1346	Crecy	Phillip VI	12000	8000	4000	Edward III	10000	9900	100	3	5
1846	Resaca de la Palma	Americans	1700	1578	122	Mexicans	5700	5083	617	3	4
1863	Brandy Station	Pleasanton	11000	10550	450	Stuart	10000	9477	523	4	8
1864	Franklin	Hood	38000	31748	6252	Schofield	32000	29644	2356	4	5
1644	Freiburg (30 YEARS WAR)	Conde	10000	5000	5000	von Mercy	15000	10000	5000	4	8
1811	Fuentes de Onoro	Massena	30000	27800	2200	Wellington	30000	28500	1500	4	12
1777	Germantown	Washington	11000	10327	673	Howe	9000	8480	520	4	12
1781	Guilford Courthouse	Cornwallis	1900	1368	532	Greene	4400	4139	261	4	6
1781	Hobkirk's Hill	Greene	1551	1417	134	Rawdon	900	642	258	4	8
1758	Hochkirch	Daun	90000	82500	7500	Frederick	37000	27000	10000	4	8
1806	Jena-AusERSTADT	Napoleon	56000	51000	5000	Hohenlohe	48000	37000	11000	4	6
1811	La Albuera	Soult	30000	22000	8000	Beresford	9000	5000	4000	4	8
1775	Lexington & Concorde	British	700	453	247	Americans	4000	3910	90	4	6
1779	Stono Ferry	Americans	1200	1054	146	British	900	711	189	4	12
1814	Brienne	Napoleon	117000	1E+05	3000	Blucher	50000	46000	4000	5	6
1757	Kolin	Frederick	35000	23000	12000	Daun	53000	45000	8000	5	5
1814	Lundy's Lane	Scott	2600	1857	743	British	3000	2357	643	5	5
1846	Palo Alto	Americans	2300	2246	54	Mexicans	6000	5650	350	5	5
1809	Raab	French	40000	37000	3000	Austria	40000	35000	5000	5	10
1861	Bull Run I	McDowell	35000	32104	2896	Bauregard	29000	27018	1982	6	10
1862	Corinth, Miss	Van Dorn	22000	19530	2470	Rosecrans	23000	20480	2520	6	12
1743	Dettingen	George II	37000	34500	2500	Charles	28000	23000	5000	6	12
1704	Donauworth	Marlborough	52000	46800	5200	Maximillian II	12000	3000	9000	6	12
1831	Grochow	von Diebitsch	100000	90000	10000	Radziwill	100000	95000	5000	6	12
1813	Hanau	Napoleon	95000	90000	5000	Wrede	43500	33500	10000	6	12
-280	Heraclea	Pyrrhus	25000	21000	4000	Laverius	35000	28000	7000	6	12
1632	Lutzen	Swedish-German	16000	14500	1500	Holy Roman	15000	12000	3000	6	10

Table D.4. Short Battles

Appendix D

Date	Battle	Attacker	Attacker Initial	Final	Losses	Defender	Initial	Final	Losses	Duration Minimum	Maximum
-190	Magnesia	Antiochus	80000	40000	40000	Scipio	40000	39700	300	6	12
1818	Maipo River	Spanish	6000	5000	1000	Chile	9000	8000	1000	6	12
1812	Maloyaroslavets	French	15000	10000	5000	Russians	20000	14000	6000	6	12
-490	Marathon	Greeks	11000	10808	192	Persians	15000	8600	6400	6	12
1759	Minden II	British	45000	38000	7000	French	60000	57000	3000	6	12
1741	Mollwitz	Prussia	30000	27500	2500	Austria	20000	15000	5000	6	12
1862	Perryville	Confederacy	44000	40855	3145	Union	39000	35304	3696	6	12
-48	Pharsalus	Pompey	40000	34000	6000	Ceasar	20000	18800	1200	6	12
1777	Saratoga	British	7000	6400	600	AMerican	3000	2681	319	6	12
1809	Talavera	French	30000	22700	7300	British	16000	10600	5400	6	12
1813	Vitoria	British	79000	71000	8000	French	66000	61000	5000	6	12
-326	Hydaspes River	Alexander	10000	9020	980	Porus	50000	38000	12000	7	9
1815	Quartre Bras	French	25000	20700	4300	Prussians	36000	31300	4700	7	7
1813	Dresden	Napoleon	70000	60000	10000	Schwarzberg	158000	1E+05	38000	8	14
1862	Fort DOnelson	Pillow	12000	10000	2000	Grant	25000	22392	2608	8	12
1870	Gravelotte	Wilhelm I	190000	2E+05	20000	Bazaine	113000	1E+05	13000	8	14
1757	Leuthen	Frederick	43000	37000	6000	Austrians	72000	66000	6000	8	12
1815	Ligny	French	77000	65500	11500	Prussians	80000	52000	28000	8	14
1813	Leipzig	Allies	355000	3E+05	52000	Napoleon	175000	1E+05	38000	9	9
1807	Friedland	Napoleon	26000	15620	10380	Bennigsen	60000	42000	18000	10	20
1813	Lutzen II	French	120000	98000	22000	Prussia	73000	53000	20000	10	10
1814	La Rotherie	Blucher	53000	47000	6000	Napoleon	40000	34000	6000	12	16
1800	Marengo	Austria	31000	21598	9402	France	23000	17165	5835	12	12
1778	Monmouth	Americans	6400	6040	360	British	13000	12642	358	12	18
1879	Rorke's Drift	Zulus	4000	3600	400	British	140	115	25	12	24
1862	Fair Oaks	Johnston	41000	34866	6134	McClelland	41000	35969	5031	16	24
1862	Pea Ridge	Confederacy	17000	16200	800	Union	11000	9616	1384	16	24
1885	Pirot	Serbs	40000	38000	2000	Bulgars	40000	37900	2100	16	24
1814	Laon	Napoleon	47000	41000	6000	Blucher	85000	81000	4000	24	48
1864	Nashville	Union	49000	46051	2949	Confederacy	31000	29500	1500	24	36
1863	Gettysburg	Lee	75000	52362	22638	Meade	88000	70316	17684	36	72
1864	Drewry's Bluff	Butler	16000	11840	4160	Bauregard	18000	15404	2596	48	72
1904	Liaoyang	Japanese	100000	76500	23500	Russians	100000	83500	16500	96	192

Table D. Short Battles

Date	Battle	Attacker	Inital	Final	Losses	Defender	Inital	Final	Losses	Duration	Winner
1914	MASURIAN LAKES	GERMAN	288600	248600	40000	RUSSIAN	273000	148000	125000	5	A
1915	WINTER BATTLE	GERMAN	250000	210000	40000	RUSSIAN	300000	200000	100000	15	A
1918	MEUSE-ARGONNE	AMERICANS	600000	470000	130000	GERMAN	380000	254000	126000	47	A
1914	TANNENBERG	GERMAN	187000	173788	13212	RUSSIAN	160000	125000	35000	4	A
1914	THE FRONTIERS	GERMANY	1200000	1000000	200000	FRENCH	1390000	1090000	300000	10	A
1915	CHAMPAGNE II	FRENCH	500000	355000	145000	GERMAN	190000	130000	60000	45	D
1914	THE MARNE	GERMAN	900000	600000	300000	FRENCH	1200000	950000	250000	6	D
1917	ARRAS	BRITISH	276000	192000	84000	GERMAN	120000	45000	75000	15	D
1917	AISNE II	FRENCH	1000000	882000	118000	GERMAN	480000	440000	40000	15	D
1918	SOMME II	GERMAN	870000	680000	190000	BRITISH	550000	350000	200000	20	S
1918	LYS	GERMAN	500000	325000	175000	BRITISH	400000	247500	152500	22	S
1914	LODZ	GERMAN	260000	200000	60000	RUSSIAN	400000	305000	95000	15	S

Table D.5. World War I Battles

APPENDIX E
AN ESSAY ON FEAR OF MATHEMATICS

FEAR OF MATHEMATICS

As we grow older and build our professional and occupational niche in life, we often find ourselves respected and trusted for some special trait or temperament. In my case, one of these is an intense interest, ability, and proclivity to use mathematics as a tool for analyzing everyday problems. There is a two part "why?" question that goes with this: Why do I do this?, and Why is it so unique?

One of the activities that I practice and enjoy is trying to teach others what I know and how to use it. Almost universally, I find that the use of math is a profound "turn-off" to the vast majority. Why is this? Why are so many people so averse to using math to analyze their daily problems and situations?

At first, I thought that people were adverse because they either lacked creativity, or had some lack of courage and will. Closer observation revealed that most people lack neither creativity nor will and courage, they ably displayed both in their professional and personal activities. In many cases, they showed larger measures of both in avoiding using math than they would have in its use. Grudgingly, I was forced to the conclusion that the respect that I had earned for my use of math derived not just from my facility with it, but also from their fearful respect for math and the rarity of my ability to use it.

I then resolved to examine why these people went to such lengths to avoid the use of math. At least part of the answer may be in answering why I do use this tool.

If I examine my childhood, and compare it to others, I find that I learned basic advanced math, algebra, and trigonometry, as a child, before I entered junior high school. I learned it on my own outside of the school environment. Actually I used some old Navy correspondence courses of my father's. By necessity, these courses were long on problems (applications) and short on theory. One result of this was that I saw no new math in school until I was a senior in high school. Another was that I learned as a child, while still adaptable, to be applications oriented and rather cavalier about whether the problem really leant itself to such analysis.

It is clear that there are two major differences between this approach and the classical method for teaching math. First, the proof of theorems is the primary component of the teaching of higher math; problem solving is minor and mechanical, not application oriented. This training may ably prepare junior mathematicians to prove new theorems, but does little to instill others with tool using capabilities, or help them view real quantities and properties in an abstract manner. Faced with this value structure, it is not amazing that the student often "turns-off".

The mathematician will decry this view, citing the necessity of technical correctness in the practice of mathematics. Perhaps surprisingly, I will agree with this, but also respond that we are talking about two different things. The mathematician is talking about the profession of mathematics where math is the environment of the profession.

I am talking about other professions where math can be a tool for manipulating that professional environment.

Consider, for a moment, the study of the calculus. For many of us, this is the first college course in mathematics. Certainly, it is often the first math course where we encounter a diet predominant in the proof of theorems. Our math training in primary and secondary schools does little to prepare us for this glut of theorems and proofs, despite some introduction to them in courses such as Analytical Geometry. Among these theorems, the most important is the Fundamental Theorem of the Calculus. Its proof has been honed by generations of mathematicians to profound elegance. Today, twenty-five years after the fact, I remember little about this theorem other than its name, but I practice its applicability as a tool every day.

Calculus is doubly difficult for most students. They must cope not only with the proofs of the theorems, but also with the mechanics for honing the mental and algebraic skills of evaluating derivatives and integrals. The former, I will contend, is not horribly difficult and is intellectually satisfying but the effort to learn it is diluted by the confusion of theorems and their proofs.

The problem is further complicated by the inattention that is devoted to helping the student learn how to make the mental association of real quantities to abstract mathematical variables and functions. This is the true skill that must be cultivated to make practical mathematicians. To make an analogy, we cannot only teach a student how to use a hammer to drive nails - we must also teach that student that nails (and pegs) can be used to build furniture if we want to teach that student how to build furniture. The same is true with mathematics. How well can we develop a furniture maker if we teach him that (i) hammers and nails exist, and (ii) furniture is made of wood. How well can we develop a practical mathematician if we only teach him that math exists but don't teach him how to apply the tools and concepts in a work-a-day context?

The classical answer is that this application framework should be taught in the professional course work. The business, scientific, and engineering disciplines make use of math in teaching their students their specific concepts and practices. Many students are lost here who, having never learned the mechanics of using math in their math classes, now must try to simultaneously learn these mechanics along with the theories and concepts of these disciplines. This divides the students' attention and effort again, so many fall by the wayside. Teachers of these courses, recognizing the lack of mathematical preparedness of their students and believing in the importance of their discipline, naturally dilute and reduce the mathematical demands of their courses to allow students to concentrate on their specific concepts and theories. In so doing, they also reduce the opportunities for the students to use math as a tool for analysis. This dilution and reduction is even more pronounced outside these disciplines. It is not amazing then that college graduates, even in the engineering and scientific disciplines, have too little appreciation of math as a tool that they can use.

I would not propose that we overnight change curricula: theorems are important because they define the structure of what can be done with math. I would propose however that we take a more pragmatic view of mathematics. I have met few people who are afraid to pick up a hammer and use it. If they were, our walls would be devoid of pictures. The use of a hammer has little to do with its manufacture for most users, who find with practice that they can use it in a meaningful manner.

The use of math is much the same. Do not fear to pick up pieces of math and use them as tools. With practice, their use will become meaningful and less clumsy. The concept of using a hammer to drive a nail in a wall to hang a picture is exactly the one that we want to apply to using math. The goal is to have a picture hanging on the wall to display information or simply to brighten our lives. The same is true with mathematics. As the use of the hammer and nail are the means to an end, so too should be our practical use of mathematics. Two abilities are needed to do this: the knowledge of practical mathematics and the ability to view real world quantities and problems as abstract mathematical terms. The first does not necessarily require formal mathematical training - reading applied mathematics books and articles, albeit probably without bogging down in the proofs of the theorems, and doing practical exercises from the texts will probably suffice - they usually do for me. The second must be cultivated by a process of familiarization with mathematical operations and results - doing exercises again!, and by reflective thought. This last is clearly difficult but is no more so than learning some craft such as counted cross stitch or woodworking.

Practical math will not solve all of our problems; it will not make the world a Utopia. It will allow us to gain new insight into our problems and make our life easier. To become practical mathematicians, we need do only three things: familiarize ourselves with the application of math, learn to see real quantities in abstract mathematical from, and control our fear of math. Of these, the last is the only truly hard one for it requires us to grow.

APPENDIX F. Gaussian Integral Approximations

In this appendix, we will consider approximations to gaussian integrals, that is, integrals over exponential functions whose arguments are quadratic. The most common source of these integrals are in the evaluation of probabilities from probability density functions.

Consider the Gaussian or Normal probability density function of zero mean,

$$p(x) = A\, e^{-\frac{x^2}{2\sigma^2}}, \tag{F-1}$$

which has the (usual) normalization condition that

$$\int_{-\infty}^{\infty} A\, e^{-\frac{x^2}{2\sigma^2}}\, dx \equiv 1. \tag{F-2}$$

We note that since the integrand is positive everywhere, we may multiple this integral by itself,

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, A^2\, e^{-\frac{x^2}{2\sigma^2}-\frac{y^2}{2\sigma^2}} = 1. \tag{F-3}$$

and introduce a change of variables from rectangular (x,y) to circular (r, θ) coordinates as

$$\begin{aligned} x &= r\cos(\theta), \\ y &= r\sin(\theta), \\ dx\,dy &= r\,dr\,d\theta. \end{aligned} \tag{F-4}$$

The normalization integral becomes, upon substitution,

$$A^2 \int_0^{2\pi} d\theta \int_0^{\infty} r\,dr\, e^{-\frac{r^2\cos^2(\theta)}{2\,\sigma^2}-\frac{r^2\sin^2(\theta)}{2\,\sigma^2}} = 1, \tag{F-5}$$

which since

$$\cos^2(\theta) + \sin^2(\theta) = 1, \tag{F-6}$$

simply becomes

$$A^2 \int_0^{2\pi} d\theta \int_0^{\infty} r\, dr\, e^{-\frac{r^2}{2\sigma^2}} = 1. \tag{F-7}$$

We note immediately that the θ integration can be trivially performed, giving us

$$2\pi A^2 \int_0^{\infty} r\, dr\, e^{-\frac{r^2}{2\sigma^2}} = 1. \tag{F-8}$$

If we now replace the argument of the exponential with a new variable z, then this integral becomes

$$2\pi \sigma^2 A^2 \int_0^{\infty} dz\, e^{-z} = 1, \tag{F-9}$$

which is also trivially integrable, and gives us

$$A = \frac{1}{\sqrt{2\pi}\,\sigma}, \tag{F-10}$$

so that the normalized Gaussian or Normal probability density function has the form,

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}}. \tag{F-11}$$

In practice, we will be commonly wanting to evaluate probability integrals of the form,

$$P(z) = \int_{-\infty}^{z} dx\, p(x). \tag{F-12}$$

If we break this integral into two parts, then we may rewrite it as

$$P(z) = \int_{-\infty}^{0} dx\, p(x) + sign(z) \int_0^{|z|} dx\, p(x), \tag{F-13}$$

and recognizing that, since the integrand is symmetric with respect to the origin, the

first right-hand-side integral has a simple value, reduce this to

$$P(z) = \frac{1}{2} + sign(z) \int_0^{|z|} dx\, p(x). \tag{F-14}$$

We might now attack the second right-hand-side integral using the change of variable technique, from rectangular to circular coordinates as before, except that we note that since the upper limit on the integration is no longer infinity, we do not map the complete plane of integration. Resultingly, the angular integral is no longer trivial.

At this point, we introduce the approximation. Since the integrand is everywhere positive, we may approximately replace the integral over the square of side $|z|$ and area z^2 with an integral over a quarter circle of equal area, thus,

$$z^2 = \frac{\pi}{4}\rho^2,$$

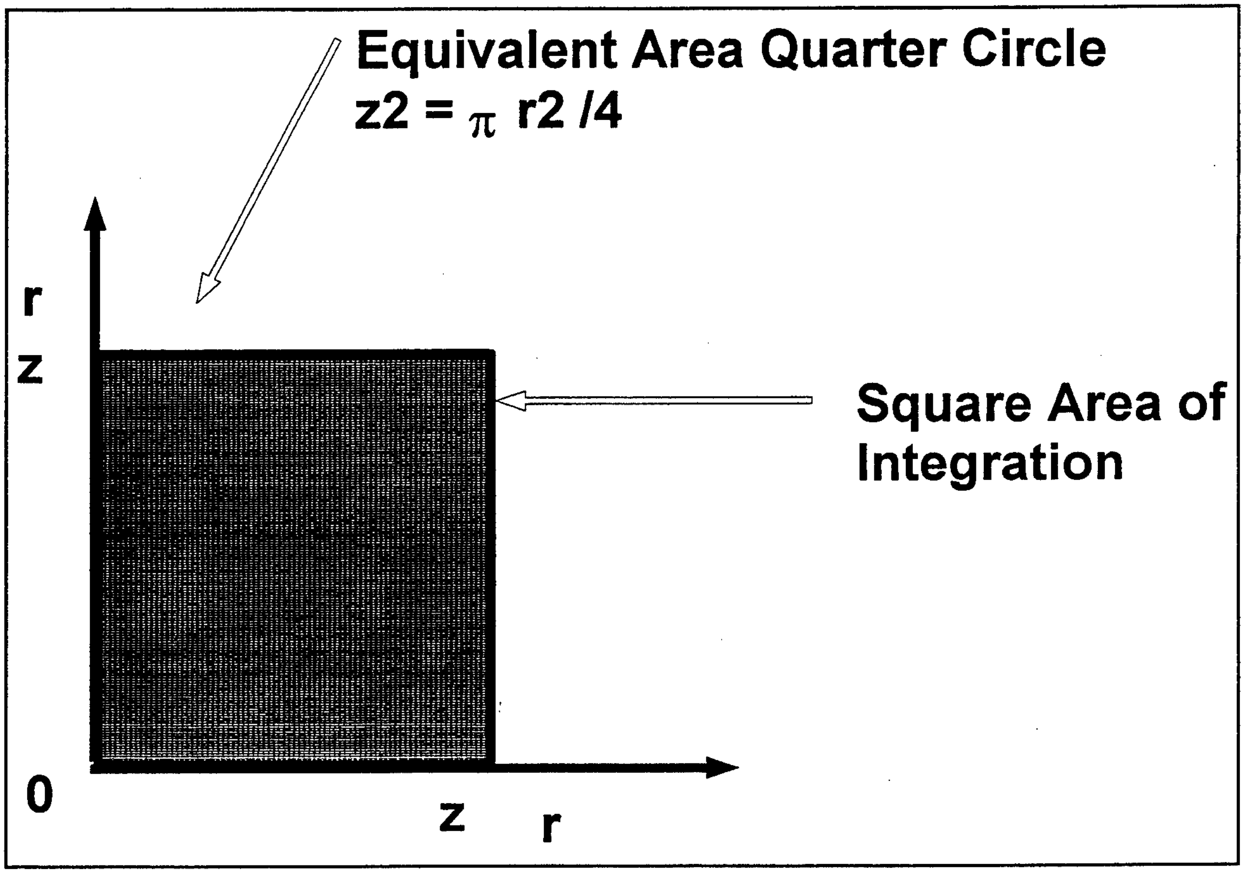


Figure F-1. Equivalent Area Quarter Circle

as shown in Figure (F-1),so that our second integral now becomes, approximately,

$$\int_0^{|z|} dz\, p(z) \approx \frac{1}{\sqrt{2\pi}\,\sigma}\sqrt{\int_0^{\frac{\pi}{2}} d\theta \int_0^{\sqrt{\frac{4}{\pi}}z} r\, dr\, e^{-\frac{r^2}{2\sigma^2}}}. \qquad \text{(F-16)}$$

As before, the θ integral, because of our assumption of the equivalent area quarter circle, is now trivial, and the r integral is exact, giving us successively,

$$\int_0^{|z|} dz\, p(z) \simeq \frac{1}{2\,\sigma}\sqrt{\int_0^{\sqrt{\frac{4}{\pi}}z} r\, dr\, e^{-\frac{r^2}{2\sigma^2}}}, \tag{F-17}$$

and

$$\int_0^{|z|} dz\, p(z) \simeq \frac{1}{2}\sqrt{1 - e^{-\frac{2\,z^2}{\pi\sigma^2}}}. \tag{F-18}$$

This then gives us the complete approximate integration of the Gaussian or Normal probability distribution function as

$$P(z) \simeq \frac{1}{2}\left(1 + sign(z)\sqrt{1 - e^{-\frac{2\,z^2}{\pi\sigma^2}}}\right). \tag{F-19}$$

This approximation has a maximum per-centum error of less than 10^{-2}. The same approximation may also be applied to the complementary probability which has the form,

$$P_c(z) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_z^{\infty} dx\, e^{-\frac{x^2}{2\sigma^2}}, \tag{F-20}$$

and gives a similar form which is just one less P(z), equation (F-19).

Of considerable more interest are the complementary integrals over the half line Gaussian probability density function, which has the similar definition,

$$p(x) = \sqrt{\frac{2}{\pi\,\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}}, \tag{F-21}$$

and the complementary integrals have the form,

$$P_c(z) = \sqrt{\frac{2}{\pi\sigma^2}} \int_z^\infty dx\, e^{-\frac{x^2}{2\sigma^2}}, \qquad \text{(F-22)}$$

so that with the equivalent are quarter circle approximation, this complementary probability has the approximate form,

$$P_c(z) \simeq e^{-\frac{z^2}{\pi\sigma^2}}, \qquad \text{(F-23)}$$

which is extremely useful in analyzing data as we shall see in the next appendix.

The same formula, divided by 2, is also valid for Equation (F-20) as long as z ≥ 0.

If we now turn to the Gaussian or Normal probability density function with non-zero mean, μ,

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \text{(F-24)}$$

then the probability integral is

$$P(z) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{z} dx\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \qquad \text{(F-25)}$$

We may introduce a change of variable,

$$y = x - \mu, \qquad \text{(F-26)}$$

so that equation (F-25) becomes

$$P(z) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{z-\mu} dy\, e^{-\frac{y^2}{2\sigma^2}}, \qquad \text{(F-27)}$$

which is identical in from to equation (F-13), and thus has the approximate solution,

$$P(z) \approx \frac{1 + sign(z - \mu)\sqrt{1 - e^{-\frac{2(z - \mu)^2}{\pi \sigma^2}}}}{2}. \tag{F-28}$$

The complementary probability is one minus this result.

Another type of probability distribution function is the log-normal probability distribution function, which has the form,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma\, x} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}}. \tag{F-29}$$

This pdf is defined on $0 \leq x \leq \infty$. The probability is

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^z \frac{dx}{x} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}}, \tag{F-30}$$

which, if we introduce the integration variable transform,

$$y = \ln(x) - \mu, \tag{F-31}$$

has the form,

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^{\ln(z) - \mu} dy\, e^{-\frac{y^2}{2\sigma^2}}. \tag{F-32}$$

This is identical in form to equation (F-27), which gives us the approximate solution,

$$P(z) \approx \frac{1 + sign(\ln(z) - \mu)\sqrt{1 - e^{-\frac{2(\ln(z) - \mu)^2}{\pi \sigma^2}}}}{2}. \tag{F-33}$$

For some distributions, such as the Negative Exponential pdf,

$$p(x) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}, \tag{F-34}$$

defined on $0 \leq x \leq \infty$, we do not need an approximation. The probability

$$P(z) = \frac{1}{\sigma} \int_0^z dx \, e^{-\frac{x}{\sigma}}, \tag{F-35}$$

is exactly integrable with form,

$$P(z) = 1 - e^{-\frac{z}{\sigma}}, \tag{F-36}$$

and has complementary probability,

$$P_c(z) = e^{-\frac{z}{\sigma}}. \tag{F-37}$$